

# OPTIMAL TRANSPORT WITH BRANCHING DISTANCE COSTS AND THE OBSTACLE PROBLEM

FABIO CAVALLETTI \*

**Abstract.** We address the Monge problem in metric spaces with a geodesic distance:  $(X, d)$  is a Polish space and  $d_N$  is a geodesic Borel distance which makes  $(X, d_N)$  a possibly branching geodesic space. We show that under some assumptions on the transference plan we can reduce the transport problem to transport problems along family of geodesics. We introduce three assumptions on a given  $d_N$ -monotone transference plan  $\pi$  which imply respectively: strongly consistency of disintegration, continuity of the conditional probabilities of the first marginal and a regularity property for the geometry of chain of transport rays. We show that this regularity is sufficient for the construction of a transport map with the same transport cost of  $\pi$ .

We apply these results to the Monge problem in  $\mathbb{R}^d$  with smooth, convex and compact obstacle obtaining the existence of an optimal map provided the first marginal is absolutely continuous w.r.t. the  $d$ -dimensional Lebesgue measure.

**Key words.** optimal transport, Monge problem, branching spaces

**AMS subject classifications.** 49J52, 28C20

**1. Introduction.** This paper concerns the Monge minimization problem in metric spaces with geodesic structure: given two Borel probability measure  $\mu, \nu \in \mathcal{P}(X)$ , where  $(X, d)$  is a Polish space, i.e. complete and separable metric space, we study the minimization of the functional

$$\mathcal{I}(T) = \int d_N(x, T(x))\mu(dy)$$

where  $T$  varies over all Borel maps  $T : X \rightarrow X$  such that  $T_\# \mu = \nu$  and  $d_N$  is a Borel distance that makes  $(X, d_N)$  a possibly branching geodesic space. We will apply the results to the obstacle problem: let  $C \subset \mathbb{R}^d$  be a convex set with  $\partial C = M$  smooth,  $(d-1)$ -dimensional compact submanifold of  $\mathbb{R}^d$ . Let  $X = (\mathbb{R}^d \setminus C) \cup M$ ,  $\mu, \nu \in \mathcal{P}(X)$  and  $d_M(x, y)$  be the infimum among all the Lipschitz curves in  $X$  connecting to  $x$  to  $y$  of the euclidean length of such curves. We will prove the existence of a solution for

$$\min_{T: T_\# \mu = \nu} \int d_M(x, T(x))\mu(dx),$$

provided  $\mu \ll \mathcal{L}^d$ .

Before describing our investigation, we present a little bit of the existing literature referring to [17] and [18] for a deeper insight into optimal transportation.

In the original formulation given by Monge in 1781 the problem was settled in  $\mathbb{R}^d$ , with the cost given by the Euclidean norm and the measures  $\mu, \nu$  were supposed to be absolutely continuous and supported on two disjoint compact sets. The original problem remained unsolved for a long time. In 1978 Sudakov [15] claimed to have a solution for any distance cost function induced by a norm: an essential ingredient in the proof was that if  $\mu \ll \mathcal{L}^d$  and  $\mathcal{L}^d$ -a.e.  $\mathbb{R}^d$  can be decomposed into convex sets of dimension  $k$ , then then the conditional probabilities are absolutely continuous with respect to the  $\mathcal{H}^k$  measure of the correct dimension. But it turns out that when  $d > 2$ ,  $0 < k < d-1$  the property claimed by Sudakov is not true. An example with  $d = 3$ ,  $k = 1$  can be found in [12].

The Euclidean case has been correctly solved only during the last decade. L. C. Evans and W. Gangbo in [9] solved the problem under the assumptions that  $\text{spt } \mu \cap \text{spt } \nu = \emptyset$ ,  $\mu, \nu \ll \mathcal{L}^d$  and their densities are Lipschitz function with compact support. The first existence results for general absolutely continuous measures  $\mu, \nu$  with compact support have been independently obtained by L. Caffarelli, M. Feldman and R.J. McCann in [6] and by N. Trudinger and X.J. Wang in [16]. Afterwards M. Feldman

---

\*SISSA, via Beirut 2, IT-34014 Trieste (ITALY), (cavallett@sisssa.it).

and R.J. McCann [10] extended the results to manifolds with geodesic cost. The case of a general norm as cost function on  $\mathbb{R}^d$ , including also the case with non strictly convex unitary ball, has been solved first in the particular case of crystalline norm by L. Ambrosio, B. Kirchheim and A. Pratelli in [2], and then in fully generality by T. Champion and L. De Pascale in [8]. The case with  $(X, d_N)$  non-branching geodesic space has been studied by S. Bianchini and the author in [4].

Concerning optimal transportation around a convex obstacle, it is worth noting that the case with transport cost  $d_M^2$  is studied in [7].

**1.1. Overview of the paper.** Let  $(X, d_N)$  be a geodesic space, not necessarily Polish. To assure that standard measure theory can be used, there exists a second distance  $d$  on  $X$  which makes  $(X, d)$  Polish and  $d_N$  Borel on  $X \times X$  with respect to the metric  $d \times d$ . We will prove that given a  $d_N$ -cyclically monotone transference plan  $\pi \in \Pi(\mu, \nu)$ , under appropriate assumptions on the first marginal and on the plan  $\pi$ , there exists an admissible map  $T : X \rightarrow X$  with the same transference cost of  $\pi$ . Since we do not require  $d_N$  to be l.s.c., the existence of an optimal transference plan is not guaranteed and our strategy doesn't rely on a possible optimality of  $\pi$ . Moreover it is worth notice that due to the lack of regularity of  $d_N$  we will not use the existence of optimal potentials  $(\phi, \psi)$ .

Our strategy to cope with the Monge problem with branching distance cost is the following:

1. reduce the problem, via Disintegration Theorem, to transportation problems in sets where, under a regularity assumption on the first marginal and on  $\pi$ , we know how to produce an optimal map;
2. show that the disintegration of the first marginal  $\mu$  on each of this sets verifies this regularity assumption;
3. find a transport map on each of these sets and piece them together.

In the easier case of  $d_N$  non-branching, given a  $d_N$ -cyclically monotone transference plan it is always possible to reduce the problem on single geodesics. The reduced problem becomes essentially one dimensional and there the precise regularity assumption is that the first marginal has no atoms (is continuous).

If  $d_N$  is a branching geodesic distance this reduction can't be done anymore and there is not another reference set where the existence of Monge minimizer is known. The reduction set will be a concatenation of more geodesics and to produce an optimal map we will need a regularity assumption also on the shape of this set.

As in the non-branching case, the reduction sets come from the class of geodesics used by a  $d_N$ -monotone plan  $\pi$ . This class can be obtained from a  $d_N$ -cyclical monotone set  $\Gamma$  on which  $\pi$  is concentrated: one can construct the set of transport rays  $R$ , the transport set  $\mathcal{T}_e$ , i.e. the set of geodesics used by  $\pi$ , and from them construct

- the set  $\mathcal{T}$  made of inner points of geodesics,
- the set  $a \cup b := \mathcal{T}_e \setminus \mathcal{T}$  of initial points  $a$  and end points  $b$ .

Since branching of geodesics is admitted,  $R$  is not a partition on  $\mathcal{T}$ . To obtain an equivalence relation we have to consider the set  $H$  of chain of transport rays: it is the set of couples  $(x, y)$  such that we can go from  $x$  to  $y$  with a finite number of transport rays such that their common points are not final or initial points.

Hence  $H$  will provide the partition of the transport set  $\mathcal{T}$  and each equivalence class,  $H(y)$  for  $y$  in the quotient space, will be a reduction set.

Even if a partition is given, the reduction to transport problems on the equivalence classes is not straightforward: a necessary and sufficient condition is that the disintegration of the measure  $\mu$  w.r.t. the partition  $H$  is strongly consistent. This is equivalent to the fact that there exists a  $\mu$ -measurable quotient map  $f : \mathcal{T} \rightarrow \mathcal{T}$  of the equivalence relation induced by the partition.

Since this partition is closely related to the geodesics of  $d_N$ , the strong consistency will follow from a topological property of the geodesic as set in  $(X, d)$  and from a metric property of  $d_N$  as a function:

- (1.a) each chain of transport rays  $H(y)$  restricted to a  $d_N$  closed ball is  $d$ -closed;
- (1.b)  $d_N(x, \cdot)$  restricted to  $H(x)$  is bounded on  $d$ -bounded sets.

Observe that these conditions on  $H$  and  $d_N$  are the direct generalization of the ones on geodesics used in [4] (continuity and local compactness) and they depend on the particular choice of the transference plan. This assumptions permit to disintegrate  $\mu$  restricted to  $\mathcal{T}$ . Hence one can write

$$\mu_{\perp \mathcal{T}} = \int \mu_y m(dy), \quad m := f_{\#} \mu, \quad \mu_y(f^{-1}(y)) = 1,$$

i.e. the conditional probabilities  $\mu_y$  are concentrated on the counterimages  $f^{-1}(y)$  (which is an equivalence class). The reduced problems are obtained by disintegrating  $\pi$  w.r.t. the partition  $H \times (X \times X)$ ,

$$\pi_{\perp \mathcal{T} \times \mathcal{T}} = \int \pi_y m(dy), \quad \nu = \int \nu_y m(dy) \quad \nu_y := (P_2)_{\#} \pi_y,$$

and considering the problems on the sets  $H(y)$  with marginals  $\mu_y, \nu_y$  and cost  $d_N$ .

To next step is study the continuity of the conditional probabilities  $\mu_y$  and whether  $\mu_{\perp \mathcal{T}_e} = \mu_{\perp \mathcal{T}}$  holds true. To pursue this aim we consider a natural operation on sets: the translation along geodesics. If  $A$  is a subset of  $\mathcal{T}$ , we denote by  $A_t$  the set translated by  $t$  in the direction determined by  $\pi$ . A rigorous definition of the translation of sets along geodesic will be given during the paper. It turns out that  $\mu(a \cup b) = 0$  and the continuity of  $\mu_y$  both depend on how the function  $t \mapsto \mu(A_t)$  behaves. Indeed assuming that:

(2) for all  $A$  Borel there exists a sequence  $\{t_n\} \subset \mathbb{R}$  and  $C > 0$  such that  $\mu(A_{t_n}) \geq C\mu(A)$  as  $t_n \rightarrow 0$ , we have the following.

**THEOREM 1.1** (Proposition 5.2 and Proposition 5.3). *If Assumption (2) holds, then  $\mu(a \cup b) = 0$  and the conditional probabilities  $\mu_y$  are continuous.*

At this level of generality we don't know how to obtain a  $d_N$ -monotone admissible map for the restricted problem even if the marginal  $\mu_y$  satisfies some regularity assumptions. Therefore we need to assume that  $H(y)$  has a particular structure:

(3) for  $m$ -a.e.  $y$ , the chain of transport rays  $H(y)$  is contained, up to set of  $\mu_y$ -measure zero, in an uncountable "increasing" family of measurable sets.

A rigorous formulation of Assumption (3) and of "increasing" will be given during the paper. If  $H(y)$  satisfies Assumption (3), then we can perform a disintegration of  $\mu_y$  with respect to the partition induced by the uncountable "increasing" family of sets. Then if the quotient measure and the marginal measures of  $\mu_y$  are continuous, we prove the existence of an optimal map between  $\mu_y$  and  $\nu_y$ .

**THEOREM 1.2** (Proposition 6.2 and Theorem 6.1). *Let  $\pi \in \Pi(\mu, \nu)$  be a  $d_N$ -monotone plan concentrated on a set  $\Gamma$ . Assume that Assumptions (1.a), (1.b), (2), (3) holds and that the quotient measure and the marginal measures of  $\mu_y$  are continuous for  $m$ -a.e.  $y$ . Then there exists an admissible map with the same transference cost of  $\pi$ .*

It follows immediately that if we also assume that  $\pi$  is optimal in the hypothesis of Theorem 1.2, then the Monge minimization problem admits a solution.

Before presenting an application of Theorem 1.2, we summarize the theoretical results. Let  $\pi \in \Pi(\mu, \nu)$  be a  $d_N$ -cyclically monotone transference plan concentrated on a set  $\Gamma$ . We consider the corresponding family of chain of transport rays and, if assumptions (1.a) and (1.b) are satisfied, we can perform, neglecting the set of initial points, a disintegration of  $\mu, \nu$  and  $\pi$  with respect to the partition induced by the chain of transport rays. Then if assumption (2) is satisfied it follows that the set of initial points is  $\mu$ -negligible and the conditional probabilities  $\mu_y$  are continuous. Since the geometry of  $H(y)$  can be wild, we need another assumption to build a  $d_N$ -monotone transference map between  $\mu_y$  and  $\nu_y$ . If  $H(y)$  satisfies assumption (3) we can perform another disintegration and, under additional regularity of the conditional probabilities of  $\mu_y$  and of the quotient measure of  $\mu_y$ , we prove the existence of a  $d_N$ -monotone transference map between  $\mu_y$  and  $\nu_y$ . Applying the same reasoning for  $m$ -a.e.  $y$  we prove the existence of a transport map  $T$  between  $\mu$  and  $\nu$  that has the same transference cost of the given  $d_N$ -cyclically monotone plan  $\pi$ .

In the last part of the paper we show an application of Theorem 1.2. Consider a hyper-surface  $M \subset \mathbb{R}^d$  that is the boundary of a convex and compact set  $C$ . Let  $X$  be the closure, in the euclidean topology, of  $\mathbb{R}^d \setminus C$  and take as cost function  $d_M$ : the minimum of the euclidean length among all Lipschitz curves in  $X$  that do not cross  $M$ . We will study the Monge minimization problem with  $C$  as convex obstacle. We will prove that if  $\mu$  is absolutely continuous w.r.t.  $\mathcal{L}^d$ , the Monge minimization problem with cost  $d_M$  admits a solution.

It is worth noting that the hypothesis of Theorem 1.2, namely Assumptions (1.a), (1.b), (2), (3), are all about the behavior of a given  $d_N$ -cyclically monotone transference plan. For the obstacle problem we will prove that any  $d_N$ -cyclically monotone transference plan satisfies this hypothesis. Therefore this concrete example provide also a confirmation of the validity of the proposed strategy and, in particular, of the non artificiality of the assumptions.

**1.2. Structure of the paper.** The paper is organized as follows.

In Section 2, we recall the mathematical tools we use in this paper. In Section 2.1 the fundamental results of projective set theory are listed. In Section 2.2 we recall the Disintegration Theorem. Next, the basic results of selection principles are listed in Section 2.3, and in Section 2.4 we define the geodesic structure  $(X, d, d_N)$  which is studied in this paper. Finally, Section 2.5 recalls some fundamental results in optimal transportation.

Section 3 shows how using only the  $d_N$ -cyclical monotonicity of a set  $\Gamma$  we can obtain a partial order relation  $G \subset X \times X$  as follows (Lemma 3.3 and Proposition 3.9):  $xGy$  iff there exists  $(w, z) \in \Gamma$  and a geodesic  $\gamma$ , passing trough  $w$  and  $z$  and with direction  $w \rightarrow z$ , such that  $x, y$  belongs to  $\gamma$  and  $\gamma^{-1}(x) \leq \gamma^{-1}(y)$ . This set  $G$  is analytic, and allows to define

- the transport rays set  $R$  (3.3),
- the transport sets  $\mathcal{T}_e, \mathcal{T}$  (with and without end points) (3.4),
- the set of initial points  $a$  and final points  $b$  (3.7).

Even if this part of Section 3 contains the same results of the first part of Section 3 of [4], for being as self contained as possible, we state this results and show their proofs again. The main difference with the non-branching case is that here  $R$  is not an equivalence relation. Therefore the approach proposed in [4] doesn't work anymore and indeed the only common part with [4] is the first part of Section 3.

To obtain an equivalence relation  $H \subset X \times X$  we have to consider the set of couples  $(x, y)$  for  $x, y \in \mathcal{T}$  such that there is a continuous path from  $x$  to  $y$ , union of a finite number of transport rays never passing through  $a \cup b$ , Definition 3.8. In Proposition 3.9 we prove that  $H$  is an equivalence relation.

Section 4 proves that the compatibility conditions (1.a) and (1.b) between  $d_N$  and  $d$  imply that the disintegration induced by  $H$  on  $\mathcal{T}$  is strongly consistent (Proposition 4.3). Using this fact we can reduce the analysis on  $H(y)$  for  $y$  in the quotient set.

In Section 5 we prove Theorem 1.1. We first introduce the operation  $A \mapsto A_t$ , the translation along geodesics (5.1), and show that  $t \mapsto \mu(A_t)$  is a  $\mathcal{A}$ -measurable function if  $A$  is analytic (Lemma 5.1). Next, we show that under the assumption

$$\mu(A) > 0 \implies \mu(A_{t_n}) \geq C\mu(A)$$

for an infinitesimal sequence  $t_n$  and  $C > 0$ , the set of initial points  $a$  is  $\mu$ -negligible (Proposition 5.2) and the conditional probabilities  $\mu_y$  are continuous.

In Section 6 we prove Theorem 1.2. First in Theorem 6.1 we prove that gluing all the  $d_N$ -cyclically monotone maps defined on  $H(y)$  we obtain a measurable transference map  $T$  from  $\mu$  to  $\nu$   $d_N$ -cyclically monotone. Then the assumption on the structure of  $\Gamma$  is stated (Assumption 3) and in Proposition 6.2 we show that on the equivalence class  $H(y)$  satisfying Assumption 3 there exists an optimal transference map  $T_y$  from  $\mu_y$  to  $\nu_y$ , provided the quotient measure and the marginal probabilities of  $\mu_y$  induced by the partition given by Assumption 3 are continuous.

Section 7 gives an application of Theorem 1.2:  $M$  is a connected smooth hyper-surface of  $\mathbb{R}^d$  that is the boundary of a convex and compact set  $C$ . Let  $X = cl(\mathbb{R}^d \setminus C)$ . The distance  $d_M$  is the minimum of

the euclidean length among all the Lipschitz curves in  $X$  (7.1). Hence  $C$  is to be intended as an obstacle for euclidean geodesics. The geodesic space  $(X, d_M)$  fits into the setting of Theorem 1.2 (Lemma 7.1 and Remark 2). If  $\mu \ll \mathcal{L}^d$  then the  $\mu$ -measure of the set of initial points is zero and the marginal  $\mu_y$  are continuous (Lemma 7.3). Finally we show in Proposition 7.4 and Proposition 7.6 that any  $d_M$ -cyclically monotone set and  $\mu$  satisfy the hypothesis of Proposition 6.2. It follows the existence of a solution for the Monge minimization problem.

**2. Preliminaries.** In this section we recall some general facts about projective classes, the Disintegration Theorem for measure, measurable selection principles, geodesic spaces and optimal transportation problems.

**2.1. Borel, projective and universally measurable sets.** The *projective class*  $\Sigma_1^1(X)$  is the family of subsets  $A$  of the Polish space  $X$  for which there exists  $Y$  Polish and  $B \in \mathcal{B}(X \times Y)$  such that  $A = P_1(B)$ . The *coprojective class*  $\Pi_1^1(X)$  is the complement in  $X$  of the class  $\Sigma_1^1(X)$ . The class  $\Sigma_1^1$  is called *the class of analytic sets*, and  $\Pi_1^1$  are the *coanalytic sets*.

The *projective class*  $\Sigma_{n+1}^1(X)$  is the family of subsets  $A$  of the Polish space  $X$  for which there exists  $Y$  Polish and  $B \in \Pi_n^1(X \times Y)$  such that  $A = P_1(B)$ . The *coprojective class*  $\Pi_{n+1}^1(X)$  is the complement in  $X$  of the class  $\Sigma_{n+1}^1$ .

If  $\Sigma_n^1, \Pi_n^1$  are the projective, coprojective pointclasses, then the following holds (Chapter 4 of [13]):

1.  $\Sigma_n^1, \Pi_n^1$  are closed under countable unions, intersections (in particular they are monotone classes);
2.  $\Sigma_n^1$  is closed w.r.t. projections,  $\Pi_n^1$  is closed w.r.t. coprojections;
3. if  $A \in \Sigma_n^1$ , then  $X \setminus A \in \Pi_n^1$ ;
4. the *ambiguous class*  $\Delta_n^1 = \Sigma_n^1 \cap \Pi_n^1$  is a  $\sigma$ -algebra and  $\Sigma_n^1 \cup \Pi_n^1 \subset \Delta_{n+1}^1$ .

We will denote by  $\mathcal{A}$  the  $\sigma$ -algebra generated by  $\Sigma_1^1$ : clearly  $\mathcal{B} = \Delta_1^1 \subset \mathcal{A} \subset \Delta_2^1$ .

We recall that a subset of  $X$  Polish is *universally measurable* if it belongs to all completed  $\sigma$ -algebras of all Borel measures on  $X$ : it can be proved that every set in  $\mathcal{A}$  is universally measurable. We say that  $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is a *Souslin function* if  $f^{-1}(t, +\infty] \in \Sigma_1^1$ .

LEMMA 2.1. *If  $f : X \rightarrow Y$  is universally measurable, then  $f^{-1}(U)$  is universally measurable if  $U$  is. See [4] for the proof.*

**2.2. Disintegration of measures.** We follow the approach of [3].

Given a measurable space  $(R, \mathcal{R})$  and a function  $r : R \rightarrow S$ , with  $S$  generic set, we can endow  $S$  with the *push forward  $\sigma$ -algebra*  $\mathcal{S}$  of  $\mathcal{R}$ :

$$Q \in \mathcal{S} \iff r^{-1}(Q) \in \mathcal{R},$$

which could be also defined as the biggest  $\sigma$ -algebra on  $S$  such that  $r$  is measurable. Moreover given a measure space  $(R, \mathcal{R}, \rho)$ , the *push forward measure*  $\eta$  is then defined as  $\eta := (r_\# \rho)$ .

Consider a probability space  $(R, \mathcal{R}, \rho)$  and its push forward measure space  $(S, \mathcal{S}, \eta)$  induced by a map  $r$ . From the above definition the map  $r$  is clearly measurable and inverse measure preserving.

DEFINITION 2.2. *A disintegration of  $\rho$  consistent with  $r$  is a map  $\rho : \mathcal{R} \times S \rightarrow [0, 1]$  such that*

1.  $\rho_s(\cdot)$  is a probability measure on  $(R, \mathcal{R})$ , for all  $s \in S$ ,
2.  $\rho(B)$  is  $\eta$ -measurable for all  $B \in \mathcal{R}$ ,

*and satisfies for all  $B \in \mathcal{R}, C \in \mathcal{S}$  the consistency condition*

$$\rho(B \cap r^{-1}(C)) = \int_C \rho_s(B) \eta(ds).$$

*A disintegration is strongly consistent with  $r$  if for all  $s$  we have  $\rho_s(r^{-1}(s)) = 1$ .*

We say that a  $\sigma$ -algebra  $\mathcal{A}$  is *essentially countably generated* with respect to a measure  $m$ , if there exists a countably generated  $\sigma$ -algebra  $\hat{\mathcal{A}}$  such that for all  $A \in \mathcal{A}$  there exists  $\hat{A} \in \hat{\mathcal{A}}$  such that  $m(A \Delta \hat{A}) = 0$ .

We recall the following version of the theorem of disintegration of measure that can be found on [11], Section 452.

**THEOREM 2.3** (Disintegration of measure). *Assume that  $(R, \mathcal{R}, \rho)$  is a countably generated probability space,  $R = \cup_{s \in S} R_s$  a decomposition of  $R$ ,  $r : R \rightarrow S$  the quotient map and  $(S, \mathcal{S}, \eta)$  the quotient measure space. Then  $\mathcal{S}$  is essentially countably generated w.r.t.  $\eta$  and there exists a unique disintegration  $s \rightarrow \rho_s$  in the following sense: if  $\rho_1, \rho_2$  are two consistent disintegration then  $\rho_{1,s}(\cdot) = \rho_{2,s}(\cdot)$  for  $\eta$ -a.e.  $s$ .*

*If  $\{S_n\}_{n \in \mathbb{N}}$  is a family essentially generating  $\mathcal{S}$  define the equivalence relation:*

$$s \sim s' \iff \{s \in S_n \iff s' \in S_n, \forall n \in \mathbb{N}\}.$$

*Denoting with  $p$  the quotient map associated to the above equivalence relation and with  $(L, \mathcal{L}, \lambda)$  the quotient measure space, the following properties hold:*

- $R_l := \cup_{s \in p^{-1}(l)} R_s = (p \circ r)^{-1}(l)$  is  $\rho$ -measurable and  $R = \cup_{l \in L} R_l$ ;
- the disintegration  $\rho = \int_L \rho_l \lambda(dl)$  satisfies  $\rho_l(R_l) = 1$ , for  $\lambda$ -a.e.  $l$ . In particular there exists a strongly consistent disintegration w.r.t.  $p \circ r$ ;
- the disintegration  $\rho = \int_S \rho_s \eta(ds)$  satisfies  $\rho_s = \rho_{p(s)}$ , for  $\eta$ -a.e.  $s$ .

In particular we will use the following corollary.

**COROLLARY 2.4.** *If  $(S, \mathcal{S}) = (X, \mathcal{B}(X))$  with  $X$  Polish space, then the disintegration is strongly consistent.*

**2.3. Selection principles.** Given a multivalued function  $F : X \rightarrow Y$ ,  $X, Y$  metric spaces, the graph of  $F$  is the set

$$\text{graph}(F) := \{(x, y) : y \in F(x)\}. \quad (2.1)$$

The *inverse image* of a set  $S \subset Y$  is defined as:

$$F^{-1}(S) := \{x \in X : F(x) \cap S \neq \emptyset\}. \quad (2.2)$$

For  $F \subset X \times Y$ , we denote also the sets

$$F_x := F \cap \{x\} \times Y, \quad F^y := F \cap X \times \{y\}. \quad (2.3)$$

In particular,  $F(x) = P_2(\text{graph}(F)_x)$ ,  $F^{-1}(y) = P_1(\text{graph}(F)^y)$ . We denote by  $F^{-1}$  the graph of the inverse function

$$F^{-1} := \{(x, y) : (y, x) \in F\}. \quad (2.4)$$

We say that  $F$  is  $\mathcal{R}$ -measurable if  $F^{-1}(B) \in \mathcal{R}$  for all  $B$  open. We say that  $F$  is *strongly Borel measurable* if inverse images of closed sets are Borel. A multivalued function is called *upper-semicontinuous* if the preimage of every closed set is closed: in particular u.s.c. maps are strongly Borel measurable.

In the following we will not distinguish between a multifunction and its graph. Note that the *domain* of  $F$  (i.e. the set  $P_1(F)$ ) is in general a subset of  $X$ . The same convention will be used for functions, in the sense that their domain may be a subset of  $X$ .

Given  $F \subset X \times Y$ , a *section*  $u$  of  $F$  is a function from  $P_1(F)$  to  $Y$  such that  $\text{graph}(u) \subset F$ . We recall the following selection principle, Theorem 5.5.2 of [13], page 198.

**THEOREM 2.5** (Von Neumann). *Let  $X$  and  $Y$  be Polish spaces,  $A \subset X \times Y$  analytic, and  $\mathcal{A}$  the  $\sigma$ -algebra generated by the analytic subsets of  $X$ . Then there is an  $\mathcal{A}$ -measurable section  $u : P_1(A) \rightarrow Y$  of  $A$ .*

A *cross-section* of the equivalence relation  $E$  is a set  $S \subset E$  such that the intersection of  $S$  with each equivalence class is a singleton. We recall that a set  $A \subset X$  is saturated for the equivalence relation  $E \subset X \times X$  if  $A = \cup_{x \in A} E(x)$ .



The next result is taken from [13], Theorem 5.2.1.

**THEOREM 2.6.** *Let  $Y$  be a Polish space,  $X$  a nonempty set, and  $\mathcal{L}$  a  $\sigma$ -algebra of subset of  $X$ . Every  $\mathcal{L}$ -measurable, closed value multifunction  $F : X \rightarrow Y$  admits an  $\mathcal{L}$ -measurable section.*

A standard corollary of the above selection principle is that if the disintegration is strongly consistent in a Polish space, then up to a saturated set of negligible measure there exists a Borel cross-section.

In particular, we will use the following corollary.

**COROLLARY 2.7.** *Let  $F \subset X \times X$  be  $\mathcal{A}$ -measurable,  $X$  Polish, such that  $F_x$  is closed and define the equivalence relation  $x \sim y \Leftrightarrow F(x) = F(y)$ . Then there exists a  $\mathcal{A}$ -section  $f : P_1(F) \rightarrow X$  such that  $(x, f(x)) \in F$  and  $f(x) = f(y)$  if  $x \sim y$ .*

*Proof.* For all open sets  $G \subset X$ , consider the sets  $F^{-1}(G) = P_1(F \cap X \times G) \in \mathcal{A}$ , and let  $\mathcal{R}$  be the  $\sigma$ -algebra generated by  $F^{-1}(G)$ . Clearly  $\mathcal{R} \subset \mathcal{A}$ .

If  $x \sim y$ , then

$$x \in F^{-1}(G) \iff y \in F^{-1}(G),$$

so that each equivalence class is contained in an atom of  $\mathcal{R}$ , and moreover by construction  $x \mapsto F(x)$  is  $\mathcal{R}$ -measurable.

We thus conclude by using Theorem 2.6 that there exists an  $\mathcal{R}$ -measurable section  $f$ : this measurability condition implies that  $f$  is constant on atoms, in particular on equivalence classes.  $\square$

**2.4. Metric setting.** In this section we refer to [5].

**DEFINITION 2.8.** *A length structure on a topological space  $X$  is a class  $\mathbf{A}$  of admissible paths, which is a subset of all continuous paths in  $X$ , together with a map  $L : \mathbf{A} \rightarrow [0, +\infty]$ : the map  $L$  is called length of path. The class  $\mathbf{A}$  satisfies the following assumptions:*

**closure under restrictions** *if  $\gamma : [a, b] \rightarrow X$  is admissible and  $a \leq c \leq d \leq b$ , then  $\gamma|_{[c, d]}$  is also admissible.*

**closure under concatenations of paths** *if  $\gamma : [a, b] \rightarrow X$  is such that its restrictions  $\gamma_1, \gamma_2$  to  $[a, c]$  and  $[c, b]$  are both admissible, then so is  $\gamma$ .*

**closure under admissible reparametrizations** *for an admissible path  $\gamma : [a, b] \rightarrow X$  and a for  $\varphi : [c, d] \rightarrow [a, b]$ ,  $\varphi \in B$ , with  $B$  class of admissible homeomorphisms that includes the linear one, the composition  $\gamma(\varphi(t))$  is also admissible.*

*The map  $L$  satisfies the following properties:*

**additivity**  $L(\gamma|_{[a, b]}) = L(\gamma|_{[a, c]}) + L(\gamma|_{[c, b]})$  for any  $c \in [a, b]$ .

**continuity**  $L(\gamma|_{[a, t]})$  is a continuous function of  $t$ .

**invariance** The length is invariant under admissible reparametrizations.

**topology** *Length structure agrees with the topology of  $X$  in the following sense: for a neighborhood  $U_x$  of a point  $x \in X$ , the length of paths connecting  $x$  with points of the complement of  $U_x$  is separated from zero:*

$$\inf \{L(\gamma) : \gamma(a) = x, \gamma(b) \in X \setminus U_x\} > 0.$$

Given a length structure, we can define a distance

$$d_N(x, y) = \inf \left\{ L(\gamma) : \gamma : [a, b] \rightarrow X, \gamma \in \mathbf{A}, \gamma(a) = x, \gamma(b) = y \right\},$$

that makes  $(X, d_N)$  a metric space (allowing  $d_N$  to be  $+\infty$ ). The metric  $d_N$  is called *intrinsic*. It follows from Proposition 2.5.9 of [5] that every admissible curve of finite length admits a constant speed parametrization, i.e.  $\gamma$  defined on  $[0, 1]$  and  $L(\gamma|_{[t, t']}) = v(t' - t)$ , with  $v$  velocity.

**DEFINITION 2.9.** *A length structure is said to be complete if for every two points  $x, y$  there exists an admissible path joining them whose length  $L(\gamma)$  is equal to  $d_N(x, y)$ . Observe that in the previous definition we do not require  $d_N(x, y) < +\infty$ .*

Intrinsic metrics associated with complete length structure are said to be *strictly intrinsic*. The metric space  $(X, d_N)$  with  $d_N$  strictly intrinsic is called a *geodesic space*. A curve whose length equals the distance between its end points is called *geodesic*.

From now on we assume the following:

1.  $(X, d)$  Polish space;
2.  $d_N : X \times X \rightarrow [0, +\infty]$  is a Borel distance;
3.  $(X, d_N)$  is a geodesic space;

Since we have two metric structures on  $X$ , we denote the quantities relating to  $d_N$  with the subscript  $N$ : for example

$$B_r(x) = \{y : d(x, y) < r\}, \quad B_{r,N}(x) = \{y : d_N(x, y) < r\}.$$

In particular we will use the notation

$$D_N(x) = \{y : d_N(x, y) < +\infty\},$$

$(\mathcal{K}, d_H)$  for the compact sets of  $(X, d)$  with the Hausdorff distance  $d_H$  and  $(\mathcal{K}_N, d_{H,N})$  for the compact sets of  $(X, d_N)$  with the Hausdorff distance  $d_{H,N}$ . We recall that  $(\mathcal{K}, d_H)$  is Polish.

LEMMA 2.10. *If  $A$  is analytic in  $(X, d)$ , then  $\{x : d_N(A, x) < \varepsilon\}$  is analytic for all  $\varepsilon > 0$ .*

*Proof.* Observe that

$$\{x : d_N(A, x) < \varepsilon\} = P_1\left(X \times A \cap \{(x, y) : d_N(x, y) < \varepsilon\}\right),$$

so that the conclusion follows from the invariance of the class  $\Sigma_1^1$  w.r.t. projections.  $\square$

In particular,  $\overline{A}^{d_N}$ , the closure of  $A$  w.r.t.  $d_N$ , is analytic if  $A$  is analytic.

**2.5. General facts about optimal transportation.** Let  $(X, \Omega, \mu)$  and  $(Y, \Sigma, \nu)$  be two probability spaces and  $c : X \times Y \rightarrow \mathbb{R}^+$  be a  $\Omega \times \Sigma$  measurable function. Consider the set of transference plans

$$\Pi(\mu, \nu) := \left\{ \pi \in \mathcal{P}(X \times Y) : (P_1)_\# \pi = \mu, (P_2)_\# \pi = \nu \right\},$$

where  $P_i(x_1, x_2) = x_i, i = 1, 2$ . Define the functional

$$\begin{aligned} \mathcal{I} & : \Pi(\mu, \nu) \longrightarrow \mathbb{R}^+ \\ \pi & \longmapsto \mathcal{I}(\pi) := \int_{X \times Y} c\pi. \end{aligned} \tag{2.5}$$

The *Monge-Kantorovich minimization problem* is the minimization of  $\mathcal{I}$  over all transference plans.

If we consider a map  $T : X \rightarrow Y$  such that  $T_\# \mu = \nu$ , the functional (2.5) becomes

$$\mathcal{I}(T) := \mathcal{I}((Id \times T)_\# \mu) = \int_X c(x, T(x)) \mu(dx).$$

The minimization problem over all  $T$  is called *Monge minimization problem*.

The Kantorovich problem admits a (pre) dual formulation: before stating it, we introduce two definitions.

DEFINITION 2.11. *A map  $\varphi : X \rightarrow \mathbb{R} \cup \{-\infty\}$  is said to be  $c$ -concave if it is not identically  $-\infty$  and there exists  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$ ,  $\psi \not\equiv -\infty$ , such that*

$$\varphi(x) = \inf_{y \in Y} \{c(x, y) - \psi(y)\}.$$



The  $c$ -transform of  $\varphi$  is the function

$$\varphi^c(y) := \inf_{x \in X} \{c(x, y) - \varphi(x)\}. \quad (2.6)$$

The  $c$ -superdifferential  $\partial^c \varphi$  of  $\varphi$  is the subset of  $X \times Y$  defined by

$$(x, y) \in \partial^c \varphi \iff c(x, y) - \varphi(x) \leq c(z, y) - \varphi(z) \quad \forall z \in X. \quad (2.7)$$

DEFINITION 2.12. A set  $\Gamma \subset X \times Y$  is said to be  $c$ -cyclically monotone if, for any  $n \in \mathbb{N}$  and for any family  $(x_1, y_1), \dots, (x_n, y_n)$  of points of  $\Gamma$ , the following inequality holds

$$\sum_{i=0}^n c(x_i, y_i) \leq \sum_{i=0}^n c(x_{i+1}, y_i),$$

with  $x_{n+1} = x_1$ . A transference plan is said to be  $c$ -cyclically monotone (or just  $c$ -monotone) if it is concentrated on a  $c$ -cyclically monotone set.

Consider the set

$$\Phi_c := \left\{ (\varphi, \psi) \in L^1(\mu) \times L^1(\nu) : \varphi(x) + \psi(y) \leq c(x, y) \right\}. \quad (2.8)$$

Define for all  $(\varphi, \psi) \in \Phi_c$  the functional

$$J(\varphi, \psi) := \int \varphi \mu + \int \psi \nu. \quad (2.9)$$

The following is a well known result (see Theorem 5.10 of [18]).

THEOREM 2.13 (Kantorovich Duality). Let  $X$  and  $Y$  be Polish spaces, let  $\mu \in \mathcal{P}(X)$  and  $\nu \in \mathcal{P}(Y)$ , and let  $c : X \times Y \rightarrow \mathbb{R}^+ \cup \{+\infty\}$  be lower semicontinuous. Then the following holds:

1. Kantorovich duality:

$$\inf_{\pi \in \Pi[\mu, \nu]} \mathcal{I}(\pi) = \sup_{(\varphi, \psi) \in \Phi_c} J(\varphi, \psi).$$

Moreover, the infimum on the left-hand side is attained and the right-hand side is also equal to

$$\sup_{(\varphi, \psi) \in \Phi_c \cap C_b} J(\varphi, \psi),$$

where  $C_b = C_b(X, \mathbb{R}) \times C_b(Y, \mathbb{R})$ .

2. If  $c$  is real valued and the optimal cost

$$C(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} I(\pi)$$

is finite, then there is a measurable  $c$ -cyclically monotone set  $\Gamma \subset X \times Y$ , closed if  $c$  is continuous, such that for any  $\pi \in \Pi(\mu, \nu)$  the following statements are equivalent:

- (a)  $\pi$  is optimal;
- (b)  $\pi$  is  $c$ -cyclically monotone;
- (c)  $\pi$  is concentrated on  $\Gamma$ ;
- (d) there exists a  $c$ -concave function  $\varphi$  such that  $\pi$ -a.s.  $\varphi(x) + \varphi^c(y) = c(x, y)$ .

3. If moreover

$$c(x, y) \leq c_X(x) + c_Y(y), \quad c_X \text{ } \mu\text{-measurable, } c_Y \text{ } \nu\text{-measurable,}$$

then there exist a couple of potentials and the optimal transference plan  $\pi$  is concentrated on the set

$$\{(x, y) \in X \times Y \mid \varphi(x) + \psi(y) = c(x, y)\}.$$

Finally if  $(c_X, c_Y) \in \mathcal{L}^1(\mu) \times \mathcal{L}^1(\nu)$  then the supremum is attained

$$\sup_{\Phi_c} J = J(\varphi, \varphi^c).$$

We recall also that if  $-c$  is Souslin, then every optimal transference plan  $\pi$  is concentrated on a  $c$ -cyclically monotone set [3].

**3. Optimal transportation in geodesic spaces.** Let  $\mu, \nu \in \mathcal{P}(X)$  and consider the transportation problem with cost  $c(x, y) = d_N(x, y)$ , and let  $\pi \in \Pi(\mu, \nu)$  be a  $d_N$ -cyclically monotone transference plan with finite cost. By inner regularity, we can assume that the optimal transference plan is concentrated on a  $\sigma$ -compact  $d_N$ -cyclically monotone set  $\Gamma \subset \{d_N(x, y) < +\infty\}$ .

Consider the set

$$\Gamma' := \left\{ (x, y) : \exists I \in \mathbb{N}_0, (w_i, z_i) \in \Gamma \text{ for } i = 0, \dots, I, z_I = y \right. \\ \left. w_{I+1} = w_0 = x, \sum_{i=0}^I d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0 \right\}. \quad (3.1)$$

In other words, we concatenate points  $(x, z), (w, y) \in \Gamma$  if they are initial and final point of a cycle with total cost 0.

LEMMA 3.1. *The following holds:*

1.  $\Gamma \subset \Gamma' \subset \{d_N(x, y) < +\infty\}$ ;
2. if  $\Gamma$  is analytic, so is  $\Gamma'$ ;
3. if  $\Gamma$  is  $d_N$ -cyclically monotone, so is  $\Gamma'$ .

*Proof.* For the first point, set  $I = 0$  and  $(w_{n,0}, z_{n,0}) = (x, y)$  for the first inclusion. If  $d_N(x, y) = +\infty$ , then  $(x, y) \notin \Gamma$  and all finite set of points in  $\Gamma$  are bounded.

For the second point, observe that

$$\Gamma' = \bigcup_{I \in \mathbb{N}_0} P_{1,2I+1}(A_I) \\ = \bigcup_{I \in \mathbb{N}_0} P_{1,2I+1} \left( \prod_{i=0}^I \Gamma \cap \left\{ \prod_{i=0}^I (w_i, z_i) : \sum_{i=0}^I d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0, w_{I+1} = w_0 \right\} \right).$$

For each  $I \in \mathbb{N}_0$ , since  $d_N$  is Borel, it follows that

$$\left\{ \prod_{i=0}^I (w_i, z_i) : \sum_{i=0}^I d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0, w_{I+1} = w_0 \right\}$$

is Borel in  $\prod_{i=0}^I (X \times X)$ , so that for  $\Gamma$  analytic each set  $A_{n,I}$  is analytic. Hence  $P_{1,2I+1}(A_I)$  is analytic, and since the class  $\Sigma_1^1$  is closed under countable unions and intersections it follows that  $\Gamma'$  is analytic.

For the third point, observe that for all  $(x_j, y_j) \in \Gamma'$ ,  $j = 0, \dots, J$ , there are  $(w_{j,i}, z_{j,i}) \in \Gamma$ ,  $i = 0, \dots, I_j$ , such that

$$d_N(x_j, y_j) + \sum_{i=0}^{I_j-1} d_N(w_{j,i+1}, z_{j,i}) - \sum_{i=0}^{I_j} d_N(w_{j,i}, z_{j,i}) = 0.$$

Hence we can write for  $x_{J+1} = x_0$ ,  $w_{j,I_j+1} = w_{j+1,0}$ ,  $w_{J+1,0} = w_{0,0}$

$$\sum_{j=0}^J d_N(x_{j+1}, y_j) - d_N(x_j, y_j) = \sum_{j=0}^J \sum_{i=0}^{I_j} d_N(w_{j,i+1}, z_{j,i}) - d_N(w_{j,i}, z_{j,i}) \geq 0,$$

using the  $d_N$ -cyclical monotonicity of  $\Gamma$ .  $\square$

DEFINITION 3.2 (Transport rays). *Define the set of oriented transport rays*

$$G := \left\{ (x, y) : \exists (w, z) \in \Gamma', d_N(w, x) + d_N(x, y) + d_N(y, z) = d_N(w, z) \right\}. \quad (3.2)$$

For  $x \in X$ , the outgoing transport rays from  $x$  is the set  $G(x)$  and the incoming transport rays in  $x$  is the set  $G^{-1}(x)$ . Define the set of transport rays as the set

$$R := G \cup G^{-1}. \quad (3.3)$$

The set  $G$  is the set of all couples of points on oriented geodesics with endpoints in  $\Gamma'$ . In  $R$  the couples are non oriented.

LEMMA 3.3. *The following holds:*

1.  $G$  is  $d_N$ -cyclically monotone;
2.  $\Gamma' \subset G \subset \{d_N(x, y) < +\infty\}$ ;
3. the sets  $G$ ,  $R := G \cup G^{-1}$  are analytic.

*Proof.* The second point follows by the definition: if  $(x, y) \in \Gamma'$ , just take  $(w, z) = (x, y)$  in the r.h.s. of (3.2).

The third point is consequence of the fact that

$$G = P_{34} \left( (\Gamma' \times X \times X) \cap \left\{ (w, z, x, y) : d_N(w, x) + d_N(x, y) + d_N(y, z) = d_N(w, z) \right\} \right),$$

and the result follows from the properties of analytic sets.

The first point follows from the following observation: if  $(x_i, y_i) \in \gamma_{[w_i, z_i]}$ , then from triangle inequality

$$\begin{aligned} d_N(x_{i+1}, y_i) - d_N(x_i, y_i) &\geq d_N(x_{i+1}, z_i) - d_N(z_i, y_i) - d_N(x_i, y_i) \\ &= d_N(x_{i+1}, z_i) - d_N(x_i, z_i) \\ &\geq d_N(w_{i+1}, z_i) - d_N(w_{i+1}, x_{i+1}) - d_N(x_i, z_i) \\ &= d_N(w_{i+1}, z_i) - d_N(w_i, z_i) + d_N(w_i, x_i) - d_N(w_{i+1}, x_{i+1}). \end{aligned}$$

Since  $(w_{n+1}, x_{n+1}) = (w_1, x_1)$ , it follows that

$$\sum_{i=1}^n d_N(x_{i+1}, y_i) - d_N(x_i, y_i) \geq \sum_{i=1}^n d_N(w_{i+1}, z_i) - d_N(w_i, z_i) \geq 0.$$

Hence the set  $G$  is  $d_N$ -cyclically monotone.  $\square$

DEFINITION 3.4. *Define the transport sets*

$$\mathcal{T} := P_1(G^{-1} \setminus \{x = y\}) \cap P_1(G \setminus \{x = y\}), \quad (3.4a)$$

$$\mathcal{T}_e := P_1(G^{-1} \setminus \{x = y\}) \cup P_1(G \setminus \{x = y\}). \quad (3.4b)$$

Since  $G$  and  $G^{-1}$  are analytic sets,  $\mathcal{T}$ ,  $\mathcal{T}_e$  are analytic. The subscript  $e$  refers to the endpoints of the geodesics: clearly we have

$$\mathcal{T}_e = P_1(R \setminus \{x = y\}). \quad (3.5)$$

The following lemma shows that we have only to study the Monge problem in  $\mathcal{T}_e$ .

LEMMA 3.5. *It holds  $\pi(\mathcal{T}_e \times \mathcal{T}_e \cup \{x = y\}) = 1$ .*

*Proof.* If  $x \in P_1(\Gamma \setminus \{x = y\})$ , then  $x \in G^{-1}(y) \setminus \{y\}$ . Similarly,  $y \in P_2(\Gamma \setminus \{x = y\})$  implies that  $y \in G(x) \setminus \{x\}$ . Hence  $\Gamma \setminus \mathcal{T}_e \times \mathcal{T}_e \subset \{x = y\}$ .  $\square$

As a consequence,  $\mu(\mathcal{T}_e) = \nu(\mathcal{T}_e)$  and any maps  $T$  such that for  $\nu \ll \mu$ ,  $T_\# \mu \ll \nu$  can be extended to a map  $T'$  such that  $\nu = T'_\# \mu$  with the same cost by setting

$$T'(x) = \begin{cases} T(x) & x \in \mathcal{T}_e \\ x & x \notin \mathcal{T}_e \end{cases} \quad (3.6)$$

DEFINITION 3.6. *Define the multivalued endpoint graphs by:*

$$a := \{(x, y) \in G^{-1} : G^{-1}(y) \setminus \{y\} = \emptyset\}, \quad (3.7a)$$

$$b := \{(x, y) \in G : G(y) \setminus \{y\} = \emptyset\}. \quad (3.7b)$$

We call  $P_2(a)$  the set of initial points and  $P_2(b)$  the set of final points.

PROPOSITION 3.7. *The following holds:*

1. *the sets*

$$a, b \subset X \times X, \quad a(A), b(A) \subset X,$$

*belong to the  $\mathcal{A}$ -class if  $A$  analytic;*

2.  *$a \cap b \cap \mathcal{T}_e \times X = \emptyset$ ;*

3.  *$a(\mathcal{T}) = a(\mathcal{T}_e)$ ,  $b(\mathcal{T}) = b(\mathcal{T}_e)$ ;*

4.  *$\mathcal{T}_e = \mathcal{T} \cup a(\mathcal{T}) \cup b(\mathcal{T})$ ,  $\mathcal{T} \cap (a(\mathcal{T}) \cup b(\mathcal{T})) = \emptyset$ .*

*Proof.* Define

$$C := \{(x, y, z) \in \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e : y \in G(x), z \in G(y)\} = (G \times X) \cap (X \times G) \cap \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e,$$

that is clearly analytic. Then

$$b = \{(x, y) \in G : y \in G(x), G(y) \setminus \{y\} = \emptyset\} = G \setminus P_{12}(C \setminus X \times \{y = z\}),$$

$$b(A) = \{y : y \in G(x), G(y) \setminus \{y\} = \emptyset, x \in A\} = P_2(G \cap A \times X) \setminus P_2(C \setminus X \times \{y = z\}).$$

A similar computation holds for  $a$ :

$$a = G^{-1} \setminus P_{23}(C \setminus \{x = y\} \times X), \quad a(A) = P_1(G_S \cap X \times A) \setminus P_1(C \setminus \{x = y\} \times X).$$

Hence  $a, b \in \mathcal{A}(X \times X)$ ,  $a(A), b(A) \in \mathcal{A}(X)$ , being the intersection of an analytic set with a coanalytic one. If  $x \in \mathcal{T}_e \setminus \mathcal{T}$ , then it follows that  $G(x) = \{x\}$  or  $G^{-1}(x) = \{x\}$  hence  $x \in a(x) \cup b(x)$ .

The other points follow easily.  $\square$

DEFINITION 3.8 (Chain of transport rays). *Define the set of chain of transport rays*

$$H := \left\{ (x, y) \in \mathcal{T}_e \times \mathcal{T}_e : \exists I \in \mathbb{N}_0, z_i \in \mathcal{T} \text{ for } 1 \leq i \leq I, \right.$$

$$(z_i, z_{i+1}) \in R, \ 0 \leq i \leq I+1, \ z_0 = x, \ z_{I+1} = y \Big\}. \quad (3.8)$$

Using similar techniques of Lemma 3.1 it can be shown that  $H$  is analytic.

PROPOSITION 3.9. *The set  $H \cap \mathcal{T} \times \mathcal{T}$  is an equivalence relation on  $\mathcal{T}$ . The set  $G$  is a partial order relation on  $\mathcal{T}_e$ .*

*Proof.* Using the definition of  $H$ , one has in  $\mathcal{T}$ :

1.  $x \in \mathcal{T}$  clearly implies that  $(x, x) \in H$ ;
2. since  $R$  is symmetric, if  $y \in H(x)$  then  $x \in H(y)$ ;
3. if  $y \in H(x)$ ,  $z \in H(y)$ ,  $x, y, z \in \mathcal{T}$ . Glue the path from  $x$  to  $y$  to the one from  $y$  to  $z$ . Since  $y \in \mathcal{T}$ ,  $z \in H(x)$ .

The second part follows similarly:

1.  $x \in \mathcal{T}_e$  implies that

$$\exists (x, y) \in (G \setminus \{x = y\}) \cup (G^{-1} \setminus \{x = y\}),$$

so that in both cases  $(x, x) \in G$ ;

2.  $(x, y), (y, z) \in G \setminus \{x = y\}$  implies by  $d_N$ -cyclical monotonicity that  $(x, z) \in G$ .

□

We finally show that we can assume that the  $\mu$ -measure of final points and the  $\nu$ -measure of the initial points are 0.

LEMMA 3.10. *The sets  $G \cap b(\mathcal{T}) \times X$ ,  $G \cap X \times a(\mathcal{T})$  is a subset of the graph of the identity map.*

*Proof.* From the definition of  $b$  one has that

$$x \in b(\mathcal{T}) \implies G(x) \setminus \{x\} = \emptyset,$$

A similar computation holds for  $a$ . □

Hence we conclude that

$$\pi(b(\mathcal{T}) \times X) = \pi(G \cap b(\mathcal{T}) \times X) = \pi(\{x = y\})$$

and following (3.6) we can assume that

$$\mu(b(\mathcal{T})) = \nu(a(\mathcal{T})) = 0.$$

**4. Partition of the transport set.** To perform a disintegration we have to assume some regularity of the support  $\Gamma$  of the transport plan  $\pi \in \Pi(\mu, \nu)$ . From now on we will assume the following:

ASSUMPTION 1. *We say that  $\Gamma$  satisfies Assumption 1 if*

- (a) *for all  $x \in \mathcal{T}$  and for all  $r > 0$  the set  $H(x) \cap \overline{B_{r,N}(x)}^{d_N}$  is  $d$ -closed;*
- (b) *for all  $x \in \mathcal{T}$  there exists  $r > 0$  such that  $d_N(x, \cdot)_{\perp H(x) \cap \overline{B_r(x)}}$  is bounded.*

Note that points (a) and (b) of Assumption 1 were already introduced at page 2. Let  $\{x_i\}_{i \in \mathbb{N}}$  be a dense sequence in  $(X, d)$ .

LEMMA 4.1. *The sets*

$$W_{ijk} := \left\{ x \in \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i) : d_N(x, \cdot)_{\perp H(x) \cap \bar{B}_{2^{-j}}(x_i)} \leq k \right\}$$

*form a countable covering of  $\mathcal{T}$  of class  $\mathcal{A}$ .*

*Proof.* We first prove the measurability. We consider separately the conditions defining  $W_{ijk}$ .

*Point 1.* The set

$$A_{ij} := \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i)$$

is clearly analytic.

*Point 2.* The set

$$D_{ijk} := \left\{ (x, y) \in H : d(x_i, y) \leq 2^{-j}, d_N(x, y) > k \right\}$$

is again analytic. We finally can write

$$W_{ijk} = A_{ij} \cap P_1(D_{ijk})^c,$$

and the fact that  $\mathcal{A}$  is a  $\sigma$ -algebra proves that  $W_{ijk} \in \mathcal{A}$ .

To show that it is a covering, notice that from (b) of Assumption 1 for all  $x \in \mathcal{T}$  there exists  $r > 0$  such that, on the set  $H(x) \cap \bar{B}_r(x)$ ,  $d_N(x, \cdot)$  is bounded. Choose  $j$  and  $i$  such that  $2^{-j-1} \leq r$  and  $d(x_i, x) \leq 2^{-j-1}$ , hence

$$\bar{B}_{2^{-j}}(x_i) \subset \bar{B}_r(x)$$

and therefore for some  $\bar{k} \in \mathbb{N}$  we obtain that  $x \in W_{ijk}$ .  $\square$

REMARK 1. Observe that  $\bar{B}_{2^{-j}}(x_i) \cap H(x)$  is closed for all  $x \in W_{ijk}$ .

Indeed take  $\{y_n\}_{n \in \mathbb{N}} \subset \bar{B}_{2^{-j}}(x_i) \cap H(x)$  with  $d(y_n, y) \rightarrow 0$  as  $n \rightarrow +\infty$ , then since  $x \in W_{ijk}$  it holds  $d_N(x, y_n) \leq k$ . By (a) of Assumption 1,  $d_N(x, y) \leq k$  and  $y \in \bar{B}_{2^{-j}}(x_i) \cap H(x)$ .

LEMMA 4.2. There exist  $\mu$ -negligible sets  $N_{ijk} \subset W_{ijk}$  such that the family of sets

$$\mathcal{T}_{ijk} = H^{-1}(W_{ijk} \setminus N_{ijk}) \cap \mathcal{T}$$

is a countable covering of  $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$  into saturated analytic sets.

*Proof.* First of all, since  $W_{ijk} \in \mathcal{A}$ , then there exists  $\mu$ -negligible set  $N_{ijk} \subset W_{ijk}$  such that  $W_{ijk} \setminus N_{ijk} \in \mathcal{B}(X)$ . Hence  $\{W_{ijk} \setminus N_{ijk}\}_{i,j,k \in \mathbb{N}}$  is a countable covering of  $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$ . It follows immediately that  $\{\mathcal{T}_{ijk}\}_{i,j,k \in \mathbb{N}}$  satisfies the lemma.  $\square$

From any analytic countable covering, we can find a countable partition into  $\mathcal{A}$ -class saturated sets by defining

$$\mathcal{Z}_m := \mathcal{T}_{i_m j_m k_m} \setminus \bigcup_{m'=1}^{m-1} \mathcal{T}_{i_{m'} j_{m'} k_{m'}}, \quad (4.1)$$

where

$$\mathbb{N} \ni m \mapsto (i_m, j_m, k_m) \in \mathbb{N}^3$$

is a bijective map. Since  $H$  is an equivalence relation on  $\mathcal{T}$ , we use this partition to prove the strong consistency.

On  $\mathcal{Z}_m$ ,  $m > 0$ , we define the closed valued map

$$\mathcal{Z}_m \ni x \mapsto F(x) := H(x) \cap \bar{B}_{2^{-j_m}}(x_{i_m}). \quad (4.2)$$

PROPOSITION 4.3. There exists a  $\mu$ -measurable cross section  $f : \mathcal{T} \rightarrow \mathcal{T}$  for the equivalence relation  $H$ .

*Proof.* First we show that  $F$  is  $\mathcal{A}$ -measurable: for  $\delta > 0$ ,

$$\begin{aligned} F^{-1}(B_\delta(y)) &= \left\{ x \in \mathcal{Z}_m : H(x) \cap B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \neq \emptyset \right\} \\ &= \mathcal{Z}_m \cap P_1 \left( H \cap (X \times B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m})) \right). \end{aligned}$$



Being the intersection of two  $\mathcal{A}$ -class sets,  $F^{-1}(B_\delta(y))$  is in  $\mathcal{A}$ . In Remark 1 we have observed that  $F$  is a closed-valued map, hence, from Lemma 5.1.4 of [13],  $\text{graph}(F)$  is  $\mathcal{A}$ -measurable.

By Corollary 2.7 there exists a  $\mathcal{A}$ -class section  $f_m : \mathcal{Z}_m \rightarrow \bar{B}_{2-j_m}(x_{i_m})$ . The proposition follows by setting  $f_{\perp \mathcal{Z}_m} = f_m$  on  $\cup_m \mathcal{Z}_m$ , and defining it arbitrarily on  $\mathcal{T} \setminus \cup_m \mathcal{Z}_m$ : the latter being  $\mu$ -negligible,  $f$  is  $\mu$ -measurable.  $\square$

Up to a  $\mu$ -negligible saturated set  $\mathcal{T}_N$ , we can assume it to have  $\sigma$ -compact range: just let  $S \subset f(\mathcal{T})$  be a  $\sigma$ -compact set where  $f_\# \mu$  is concentrated, and set

$$\mathcal{T}_S := H^{-1}(S) \cap \mathcal{T}, \quad \mathcal{T}_N := \mathcal{T} \setminus \mathcal{T}_S, \quad \mu(\mathcal{T}_N) = 0. \quad (4.3)$$

Hence we have a measurable cross-section

$$\mathcal{S} := S \cup f(\mathcal{T}_N) = (\text{Borel}) \cup (f(\mu\text{-negligible})).$$

Hence Disintegration Theorem 2.3 yields

$$\mu_{\perp \mathcal{T}} = \int_S \mu_y m(dy), \quad m = f_\# \mu_{\perp \mathcal{T}}, \quad \mu_y \in \mathcal{P}(H(y)) \quad (4.4)$$

and the disintegration is strongly consistent since the quotient map  $f : \mathcal{T} \rightarrow \mathcal{T}$  is  $\mu$ -measurable and  $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$  is countably generated.

Observe that  $H$  induces an equivalence relation also on  $\mathcal{T} \times X \cap \Gamma$  where the equivalence classes are  $H(y) \cap \mathcal{T} \times X$  and the quotient map is the  $f$  of Proposition 4.3. Hence

$$\pi_{\perp \mathcal{T} \times X \cap \Gamma} = \int_S \pi_y m_\pi(dy), \quad m_\pi = f_\# \pi_{\perp \mathcal{T} \times X \cap \Gamma}, \quad \pi_y \in \mathcal{P}(H(y) \cap \mathcal{T} \times X). \quad (4.5)$$

Observe that  $m = m_\pi$ .

**5. Regularity of the disintegration.** In this Section we consider the translation of Borel sets by the optimal geodesic flow, we introduce the fundamental regularity assumption (Assumption 2) on the measure  $\mu$  and we show that an immediate consequence is that the set of initial points is negligible and consequently we obtain a disintegration of  $\mu$  on the whole space. A second consequence is that the disintegration of  $\mu$  w.r.t. the  $H$  has continuous conditions probabilities.

**5.1. Evolution of Borel sets.** Let  $A \subset \mathcal{T}_e$  be an analytic set and define for  $t \in \mathbb{R}$  the  $t$ -evolution  $A_t$  of  $A$  by:

$$A_t := \begin{cases} P_2\{(x, y) \in G \cap A \times X : d_N(x, y) = t\} & t \geq 0 \\ P_2\{(x, y) \in G^{-1} \cap A \times X : d_N(x, y) = t\} & t < 0. \end{cases} \quad (5.1)$$

It is clear from the definition that if  $A$  is analytic, also  $A_t$  is analytic. We can show that  $t \mapsto \mu(A_t)$  is measurable.

LEMMA 5.1. *Let  $A$  be analytic. The function  $t \mapsto \mu(A_t)$  is  $\mathcal{A}$ -measurable for  $t \in \mathbb{R}$ .*

*Proof.* We divide the proof in three steps.

*Step 1.* Define the subset of  $X \times \mathbb{R}$

$$\hat{A} := \{(x, t) : x \in A_t\}.$$

Note that

$$\hat{A} = P_{13} \left\{ (x, y, t) \in X \times X \times \mathbb{R}^+ : (x, y) \in G \cap A \times X, d_N(x, y) = t \right\}$$

$$\cup P_{13} \left\{ (x, y, t) \in X \times X \times \mathbb{R}^- : (x, y) \in G^{-1} \cap A \times X, d_N(x, y) = -t \right\},$$

hence it is analytic. Clearly  $A_t = \hat{A}(t)$ .

*Step 2.* Define the closed set in  $\mathcal{P}(X \times [0, 1])$

$$\Pi(\mu) := \{ \pi \in \mathcal{P}(X \times [0, 1]) : (P_1)_\#(\pi) = \mu \}$$

and let  $B \subset X \times \mathbb{R} \times [0, 1]$  be a Borel set such that  $P_{12}(B) = \hat{A}$ .

Consider the function

$$\mathbb{R} \times \Pi(\mu) \ni (t, \pi) \mapsto \pi(B(t)).$$

A slight modification of Lemma 4.12 in [3] shows that this function is Borel.

*Step 3.* Since supremum of Borel function are  $\mathcal{A}$ -measurable, pag. 134 of [13], the proof is concluded once we show that

$$\mu(A_t) = \mu(\hat{A}(t)) = \sup_{\pi \in \Pi(\mu)} \pi(B(t)).$$

Since  $\hat{A}(t) \times [0, 1] \supset B(t)$

$$\mu(\hat{A}(t)) = \pi(\hat{A}(t) \times [0, 1]) \geq \pi(B(t)).$$

On the other hand from Theorem 2.5, there exists an  $\mathcal{A}$ -measurable section of the analytic set  $B(t)$ , so we have  $u : \hat{A}(t) \rightarrow B(t)$ . Clearly for  $\pi_u = (\mathbb{I}, u)_\#(\mu)$  it holds  $\pi_u(B(t)) = \mu(\hat{A}(t))$ .

□

The next assumption is the fundamental assumption of the paper.

**ASSUMPTION 2** (Non-degeneracy assumption). *The measure  $\mu$  satisfies Assumption 2 if for each analytic set  $A \subset \mathcal{T}_e$  there exists a sequence  $\{t_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$  and a strictly positive constant  $C$  such that  $t_n \rightarrow 0$  as  $n \rightarrow +\infty$  and  $\mu(A_{t_n}) \geq C\mu(A)$  for every  $n \in \mathbb{N}$ .* Note that Assumption 2 was already introduced at page 3. Clearly it is enough to verify Assumption 2 for  $A$  compact set. An immediate consequence of the Assumption 2 is that the measure  $\mu$  is concentrated on  $\mathcal{T}$ .

**PROPOSITION 5.2.** *If  $\mu$  satisfies Assumption 2 then*

$$\mu(\mathcal{T}_e \setminus \mathcal{T}) = 0.$$

*Proof.* Let  $A = \mathcal{T}_e \setminus \mathcal{T}$ . Suppose by contradiction  $\mu(A) > 0$ . By the inner regularity there exists  $\hat{A} \subset A$  closed with  $\mu(\hat{A}) > 0$ . By Assumption 2 there exist  $C > 0$  and  $\{t_n\}_{n \in \mathbb{N}}$  converging to 0 such that  $\mu(\hat{A}_{t_n}) \geq C\mu(\hat{A})$ .

Define  $\hat{A}^\varepsilon := \{x : d_N(\hat{A}, x) < \varepsilon\}$ . Since  $\hat{A} \subset A$ , for all  $n \in \mathbb{N}$  it holds  $\hat{A}_{t_n} \cap A = \emptyset$ . Moreover for  $t_n \leq \varepsilon$  we have  $\hat{A}^\varepsilon \supset \hat{A}_{t_n}$ . So we have

$$\mu(\hat{A}) = \lim_{\varepsilon \rightarrow 0} \mu(\hat{A}^\varepsilon) \geq \mu(\hat{A}) + \mu(\hat{A}_{t_n}) \geq (1 + C)\mu(\hat{A}),$$

that gives the contradiction. □

Once we know that  $\mu(\mathcal{T}) = 1$ , we can use the Disintegration Theorem 2.3 to write

$$\mu = \int_S \mu_y m(dy), \quad m = f_\# \mu, \quad \mu_y \in \mathcal{P}(H(y)). \quad (5.2)$$

The disintegration is strongly consistent since the quotient map  $f : \mathcal{T} \rightarrow \mathcal{T}$  is  $\mu$ -measurable and  $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$  is countably generated.

The second consequence of Assumption 2 is that  $\mu_y$  is continuous, i.e.  $\mu_y(\{x\}) = 0$  for all  $x \in X$ .

**PROPOSITION 5.3.** *If  $\mu$  satisfies Assumption 2 then the conditional probabilities  $\mu_y$  are continuous for  $m$ -a.e.  $y \in S$ .*

*Proof.* From the regularity of the disintegration and the fact that  $m(S) = 1$ , we can assume that the map  $y \mapsto \mu_y$  is weakly continuous on a compact set  $K \subset S$  of comeasure  $< \varepsilon$ . It is enough to prove the proposition on  $K$ .

*Step 1.* From the continuity of  $K \ni y \mapsto \mu_y \in \mathcal{P}(X)$  w.r.t. the weak topology, it follows that the map

$$y \mapsto A(y) := \{x \in H(y) : \mu_y(\{x\}) > 0\} = \cup_n \{x \in H(y) : \mu_y(\{x\}) \geq 2^{-n}\}$$

is  $\sigma$ -closed: in fact, if  $(y_m, x_m) \rightarrow (y, x)$  and  $\mu_{y_m}(\{x_m\}) \geq 2^{-n}$ , then  $\mu_y(\{x\}) \geq 2^{-n}$  by u.s.c. on compact sets. Hence  $A$  is Borel, where  $A = \{(y, A(y)) : y \in K\}$ .

*Step 2.* The claim is equivalent to  $\mu(P_2(A)) = 0$ . Suppose by contradiction  $\mu(P_2(A)) > 0$ . By Lusin Theorem (Theorem 5.8.11 of [13])  $A$  is the countable union of Borel graphs,  $A = \cup_n A_n$ . Therefore we can take a Borel selection of  $A$  just considering one of the Borel graphs, say  $\hat{A}$ . Since at least one of  $P_2(A_n)$  must have positive  $\mu$ -measure, we can assume  $\mu(P_2(\hat{A})) > 0$ .

By Assumption 2  $\mu((P_2(\hat{A}))_{t_n}) \geq C\mu(P_2(\hat{A}))$  for some  $C > 0$  and  $t_n \rightarrow 0$ . Since  $\hat{A}$  is a Borel graph, for every  $y \in P_1(\hat{A})$  the set  $P_2(\{y\} \times X \cap \hat{A})$  is a singleton. Hence  $(P_2(\hat{A}))_{t_n} \cap (P_2(\hat{A}))_{t_m} = \emptyset$ . We have a contradiction with the fact that the measure is finite.  $\square$

**6. Solution to the Monge problem.** Throughout the section we assume  $\mu$  to satisfy Assumption 2. It follows from Disintegration Theorem 2.3, Proposition 5.2 and Proposition 5.3 that

$$\mu = \int \mu_y m(dy), \quad \pi = \int \pi_y m(dy), \quad \mu_y \text{ continuous, } (P_1)_\# \pi_y = \mu_y,$$

where  $m = f_\# \mu$  and  $\mu_y \in \mathcal{P}(H(y))$ . We write moreover

$$\nu = \int \nu_y m(dy) = \int (P_2)_\# \pi_y m(dy).$$

Note that  $\pi_y \in \Pi(\mu_y, \nu_y)$  is  $d_N$ -cyclically monotone and (since  $d_{N \perp H(y) \times H(y)}$  is finite and, from point (a) of Assumption 1, lower semi-continuous) optimal for  $m$ -a.e.  $y$ . If  $\nu(\mathcal{T}) = 1$ , then the above formula is the disintegration of  $\nu$  w.r.t.  $H$ .

**THEOREM 6.1.** *Assume that for all  $y \in S$  there exists an optimal map  $T_y$  from  $\mu_y$  to  $\nu_y$ . Then there exists a  $\mu$ -measurable map  $T : X \rightarrow X$  such that*

$$\int d_N(x, T(x)) \mu(dx) = \int d_N(x, z) \pi(dxdz), \quad T_\# \mu = \nu.$$

Recall  $S \subset \mathcal{T}$  introduced in (4.3).

*Proof.* The idea is to use Theorem 2.5.

*Step 1.* Let  $\mathbf{T} \subset S \times \mathcal{P}(X^2)$  be the set: for  $y \in S$ ,  $\mathbf{T}_y$  is the family of optimal transference plans in  $\Pi(\tilde{\mu}_y, \tilde{\nu}_y)$  concentrated on a graph,

$$\mathbf{T} = \left\{ (y, \pi) \in S \times \mathcal{P}(X^2) : \pi \in \Pi(\mu_y, \nu_y) \text{ optimal, } \exists T : X \rightarrow X, \pi(\text{graph}(T)) = 1 \right\}.$$

where for optimal in  $\Pi(\mu_y, \nu_y)$  we mean

$$\int d_N \pi = \min_{\pi \in \Pi(\mu_y, \nu_y)} \int d_N \pi.$$

Note that, since  $\pi$  is a Borel measure, in the definition of  $\mathbf{T}$ ,  $T$  can be taken Borel. Moreover the  $y$  section  $\mathbf{T}_y = \mathbf{T} \cap \{y\} \times \mathcal{P}(X^2)$  is not empty.

*Step 2.* Since the projection is a continuous map, then the set

$$\tilde{\Pi} = \left\{ (y, \pi) : (P_1)_\# \pi = \mu_y, (P_2)_\# \pi = \nu_y \right\}$$

is a Borel subsets of  $S \times \mathcal{P}(X^2)$ : in fact it is the counter-image of the Borel set  $\text{graph}((\mu_y, \nu_y)) \subset S \times \mathcal{P}(X)^2$  w.r.t. the weakly continuous map  $(y, \pi) \mapsto (y, (P_1)_\# \pi, (P_2)_\# \pi)$ .

Define the Borel function

$$S \times \mathcal{P}(X^2) \ni (y, \pi) \mapsto f(y, \pi) := \begin{cases} \int d_N \pi & \pi \in \Pi(\mu_y, \nu_y) \\ +\infty & \text{otherwise} \end{cases}$$

It follows that  $y \mapsto g(y) := \inf_\pi f(y, \pi)$  is an  $\mathcal{A}$ -function: we can redefine it on a  $m$ -negligible set to make it Borel, where  $m$  is the quotient measure of  $\mu$ . Hence the set

$$\tilde{\Pi}^{\text{opt}} = \left\{ (y, \pi) : \pi \in \tilde{\Pi}(\mu_y, \nu_y), \int d_N \pi \leq g(y) \right\} = \tilde{\Pi} \cap \left\{ (y, \pi) : \int d_N \pi \leq g(y) \right\}$$

is Borel.

*Step 3.* Now we show that the set of  $\pi \in \mathcal{P}(X^2)$  concentrated on a graph is analytic. By Borel Isomorphism Theorem, see [13] page 99, it is enough to prove the same statement for  $\pi \in \mathcal{P}([0, 1])^2$ . Consider the function

$$\mathcal{P}([0, 1]^2) \times C_b([0, 1], [0, 1]) \ni (\pi, \phi) \mapsto h(\pi, \phi) := \pi(\text{graph}(\phi)) \in [0, 1].$$

Since  $\text{graph}(\phi)$  is compact,  $h$  is u.s.c.. Hence the set  $B^n = h^{-1}([1 - 2^{-n}, 1])$  is closed, so that

$$\mathcal{T} = \bigcap_n P_1(B^n) = \left\{ \pi : \forall \varepsilon > 0 \exists \phi_\varepsilon, \pi(\phi_\varepsilon) > 1 - \varepsilon \right\}$$

is an analytic set. It is easy to prove that  $\pi \in \mathcal{T}$  iff  $\pi$  is concentrated on a graph.

*Step 4.* It follows that

$$\mathbf{T} = S \times \mathcal{T} \cap \tilde{\Pi}^{\text{opt}}$$

is analytic and by Theorem 2.5 there exists a  $m$ -measurable selection  $y \mapsto \pi_y \in \mathbf{T}_y$ . It is fairly easy to prove that  $\int \pi_y m(dy)$  is concentrated on a graph, has the same transference cost of  $\pi$  and belongs to  $\Pi(\mu, \nu)$ .  $\square$

It follows from Theorem 6.1 that it is enough to solve for each  $y \in S$  the Monge minimization problem with marginal  $\mu_y$  and  $\nu_y$  on the set  $H(y)$ . In order to solve it, we introduce an assumption on the geometry of the set  $H(y)$ .

**ASSUMPTION 3.** *For a given  $y \in S$ ,  $H(y)$  satisfies Assumption 3 if there exist two families of disjoint  $\mathcal{A}$ -measurable sets  $\{K_t\}_{t \in [0, 1]}$  and  $\{Q_s\}_{s \in [0, 1]}$  such that*

- $\mu_y(H(y) \setminus \cup_{t \in [0, 1]} K_t) = \nu_y(H(y) \setminus \cup_{s \in [0, 1]} Q_s) = 0$ ;
- the associated quotient maps  $\varphi_K$  and  $\varphi_Q$  are respectively  $\mu_y$ -measurable and  $\nu_y$ -measurable;
- for  $t \leq s$ ,  $K_t \times Q_s \subset G$ .

Note that Assumption 3 was already introduced at page 3. In the measurability condition of Assumption 3, the set  $[0, 1]$  is equipped with the Borel  $\sigma$ -algebra  $\mathcal{B}([0, 1])$ . If  $H(y)$  satisfies Assumption 3 we can disintegrate the marginal measures  $\mu_y$  and  $\nu_y$  respectively w.r.t. the family  $\{K_t\}$  and  $\{Q_s\}$ :

$$\mu_y = \int \mu_{y,t} m_{\mu_y}(dt), \quad \nu_y = \int \nu_{y,t} m_{\nu_y}(dt)$$

where  $m_{\mu_y} = \varphi_K \# \mu_y$ ,  $m_{\nu_y} = \varphi_Q \# \nu_y$  and the disintegrations are strongly consistent.

PROPOSITION 6.2. *Suppose that  $H(y)$  satisfies Assumption 3 and that the following conditions hold true:*

- $m_{\mu_y}$  is continuous;
- $\mu_{y,t}$  is continuous for  $m_{\mu_y}$ -a.e.  $t \in [0, 1]$ ;
- $m_{\mu_y}([0, t]) \geq m_{\nu_y}([0, t])$  for  $m_{\mu_y}$ -a.e.  $t \in [0, 1]$ .

*Then there exists a  $d_N$ -cyclically monotone  $\mu_y$ -measurable map  $T_y$  such that  $T_y \# \mu_y = \nu_y$  and*

$$\int d_N(x, T_y(x)) \mu_y(dx) = \int d_N(x, z) \pi_y(dxdz).$$

*Proof. Step 1.* Since  $m_{\mu_y}$  is continuous and  $m_{\mu_y}([0, t]) \geq m_{\nu_y}([0, t])$ , there exists an increasing map  $\psi : [0, 1] \rightarrow [0, 1]$  such that  $\psi \# m_{\mu_y} = m_{\nu_y}$ .

Moreover, since for  $m_{\mu_y}$ -a.e.  $t \in [0, 1]$   $\mu_{y,t}$  is continuous, there exists a Borel map  $T_t : K_t \rightarrow Q_{\psi(t)}$  such that  $T_t \# \mu_{y,t} = \nu_{y,\psi(t)}$  for  $m_{\mu_y}$ -a.e.  $t \in [0, 1]$ . Since  $\psi(t) \geq t$  the map  $T_t$  is  $d_N$ -cyclically monotone, hence optimal between  $\mu_{y,t}$  and  $\nu_{y,t}$ .

*Step 2.* Reasoning as in the proof of Theorem 6.1, one can prove the existence of a  $\mu_y$ -measurable map  $T_y : H(y) \rightarrow H(y)$  that is the gluing of all the maps  $T_t$  constructed in *Step 1*. Hence there exists a  $\mu_y$ -measurable map  $T_y : H(y) \rightarrow H(y)$  such that  $T_y \# \mu_{y,t} = \nu_{y,\psi(t)}$ . It follows from Assumption 3 that  $\text{graph}(T_y) \subset G$ , hence  $T_y$  is  $d_N$ -cyclically monotone and

$$T_y \# \mu_y = \int T_y \# \mu_{y,t} m_{\mu_y}(dt) = \int \nu_{y,\psi(t)} m_{\mu_y}(dt) = \int \nu_{y,t} (\psi \# m_{\mu_y})(dt) = \nu_y.$$

□

The next corollary follows straightforwardly and it sums up all the results.

COROLLARY 6.3. *Let  $\pi \in \Pi(\mu, \nu)$  be concentrated on a  $d_N$ -cyclically monotone set  $\Gamma$  satisfying Assumption 1. Assume that  $\mu$  satisfies Assumption 2 and for  $m$ -a.e.  $y \in S$  the set  $H(y)$  satisfies Assumption 3. If for  $m$ -a.e.  $y \in S$  the hypothesis of Proposition 6.2 are verified, then there exists an Borel map  $T : X \rightarrow X$  such that*

$$\int d_N(x, T(x)) \mu(dx) = \int d_N(x, z) \pi(dxdz), \quad T \# \mu = \nu.$$

*If  $\pi$  is also optimal, then  $T$  solves the Monge minimization problem.*

Let us summarize the theoretical results obtained so far. Let  $\pi \in \Pi(\mu, \nu)$  be a  $d_N$ -cyclically monotone transference plan concentrated on a set  $\Gamma$ . Consider the corresponding family of chain of transport rays and assume that  $\Gamma$  satisfies Assumption 1. Then the partition induced by  $H$  permits to obtain a strongly consistent disintegration formula of  $\mu, \nu$  and  $\pi$  holds. If  $\mu$  satisfies Assumption 2 then the set of initial points is  $\mu$ -negligible and the conditional probabilities  $\mu_y$  are continuous.

Since the geometry of  $H(y)$  can be wild, we need another assumption to build a  $d_N$ -monotone transference map between  $\mu_y$  and  $\nu_y$ . If  $H(y)$  satisfies Assumption 3 we can perform another disintegration and, under additional regularity of the conditional probabilities of  $\mu_y$  and of the quotient measure of  $\mu_y$ , we prove the existence of a  $d_N$ -monotone transference map between  $\mu_y$  and  $\nu_y$ . Applying the same reasoning for  $m$ -a.e.  $y$  we prove the existence of a transport map  $T$  between  $\mu$  and  $\nu$  that has the same transference cost of the given  $d_N$ -cyclically monotone plan  $\pi$ .

**6.1. Example.** We conclude this Section with the analysis of a particular case in which the set  $H(y)$  satisfies Assumption 3. The hypothesis of Proposition 6.2 and Assumption 3 were partially inspired by this example. What follows will be useful in the next Section, however, since it is not only related to what will be proved in Section 7, we have decided to present it here.

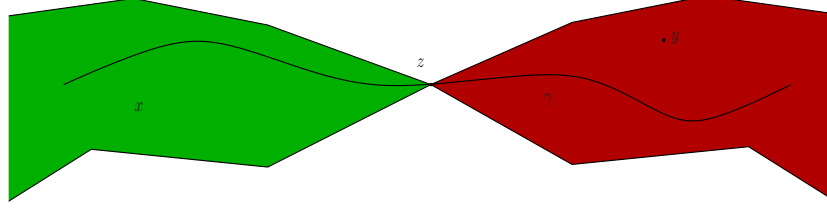


Fig. 6.1: The hourglass set  $K(z)$ .

Fix the following notation: a continuous curve  $\gamma : [0, 1] \rightarrow X$  is *increasing* if for  $t, s \in [0, 1]$

$$t \leq s \implies (\gamma(t), \gamma(s)) \in G$$

DEFINITION 6.4 (Hourglass sets). For  $z \in X$  define the hourglass set

$$K(z) := \left\{ (x, y) \in X \times X : (x, z), (z, y) \in G \right\}.$$

Assume that there exists an increasing curve  $\gamma$  such that

$$H(y) \times X \cap \Gamma \subset \bigcup_{t \in [0, 1]} K(\gamma(t)) \cap \Gamma.$$

Note that this assumption is equivalent to request that on each chain of transport rays the branching structures can appear only along an increasing curve  $\gamma$ .

Then  $H(y)$  satisfies Assumption 3. Indeed first notice that  $K(z)$  is analytic, then define the family of sets

$$K_t := G^{-1}(\gamma(t)) \setminus \bigcap_{s < t} G^{-1}(\gamma(s)), \quad Q_t := G(\gamma(t)) \setminus \bigcap_{t < s} G(\gamma(s)).$$

Since  $\gamma$  is increasing,  $K_t$  and  $Q_s$  are  $\mathcal{A}$ -measurable and the quotient maps are  $\mathcal{A}$ -measurable: let  $[a, b] \subset [0, 1]$

$$\varphi_K^{-1}([a, b]) = \bigcup_{t \in [a, b]} K_t = G^{-1}(\gamma(b)) \setminus G^{-1}(\gamma(a)) \cup K_a \in \mathcal{A}$$

and the same calculation holds true for  $\varphi_Q$ . From the increasing property of  $\gamma$  it follows that  $K_t \times Q_s \in G$  for every  $0 \leq t \leq s \leq 1$ . Again from the increasing property of  $\gamma$  it follows that  $m_{\mu_y}([0, t]) \geq m_{\nu_y}([0, t])$ .

**7. An application.** Throughout this section  $|\cdot|$  will be the euclidean distance of  $\mathbb{R}^d$ .

Let  $C \subset \mathbb{R}^d$  be an open convex set such that  $M := \partial C$  is a smooth compact sub-manifold of  $\mathbb{R}^d$  of dimension  $d - 1$ . Let  $X := \mathbb{R}^d \setminus C$ . Clearly  $X$  endowed with the euclidean topology is a Polish space.

Consider the following geodesic distance:  $d_M : X \times X \rightarrow [0, +\infty]$ :

$$d_M(x, y) := \inf \{ L(\gamma) : \gamma \in \text{Lip}([0, 1], X), \gamma(0) = x, \gamma(1) = y \}, \quad (7.1)$$

where  $L$  is the standard euclidean arc-length:  $L(\gamma) = \int |\dot{\gamma}|$ . Hence  $M$  can be seen as an obstacle for geodesics connecting points in  $X$ . Note that any minimizing sequence has uniformly bounded Lipschitz



constant, therefore in the definition of  $d_M$  we can substitute  $\inf$  with  $\min$ . Hence  $d_M$  is a geodesic distance on  $X$ .

We will show that given  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$  with  $\mu \ll \mathcal{L}^d$ , the Monge minimization problem with geodesic cost  $d_M$  admits a solution.

From now on we will assume that  $\mu \ll \mathcal{L}^d$  and all the sets and structures introduced during the paper will be referred to this Monge problem.

The strategy to solve the Monge minimization problem is the one used in Section 6: build an optimal map on each equivalence class  $H(y)$  and then use Theorem 6.1. To prove the existence these optimal maps we will show that the geometry of the chain of transport rays  $H(y)$  is the one presented in Example 6.1 and that the hypothesis of Proposition 6.2 are satisfied.

LEMMA 7.1. *The distance  $d_M$  is a continuous map.*

*Proof. Step 1.* Let  $\{x_n\}_{n \in \mathbb{N}}, \{y_n\}_{n \in \mathbb{N}} \in X$  such that  $|x_n - x| \rightarrow 0, |y_n - y| \rightarrow 0$ . Since the boundary of  $X$  is a smooth manifold, for every  $n \in \mathbb{N}$  there exist curves  $\gamma_{1,n}, \gamma_{2,n} \in \text{Lip}([0, 1], X)$  such that

- $\gamma_{1,n}(0) = x, \gamma_{1,n}(1) = x_n$ ;
- $\gamma_{2,n}(0) = y, \gamma_{2,n}(1) = y_n$ ;
- $L(\gamma_{i,n}) \rightarrow 0$  as  $n \rightarrow +\infty$ , for  $i = 1, 2$ .

Consider  $\gamma_n \in \text{Lip}([0, 1], X)$  such that  $\gamma_n(0) = x_n, \gamma_n(1) = y_n$  and  $L(\gamma_n) \leq d_M(x_n, y_n) + 2^{-n}$ . Gluing  $\gamma_{1,n}$  and  $\gamma_{2,n}$  to  $\gamma_n$  it follows

$$d_M(x, y) \leq d_M(x_n, y_n) + 2^{-n} + L(\gamma_{1,n}) + L(\gamma_{2,n}).$$

Hence  $d_M$  is l.s.c..

*Step 2.* Taking a minimizing sequence of admissible curves for  $d_M(x, y)$  and gluing them with  $\gamma_{i,n}$  as in *Step 1.*, it is fairly easy to prove that  $d_M$  is u.s.c. and therefore continuous.  $\square$

As a corollary we have the existence of an optimal transference plan  $\pi$ . Hence from now on  $\pi$  will be an optimal transference plan and all the structures defined during the paper starting from a generic  $d_N$ -cyclically monotone plan, are referred to it. Moreover there exists  $\varphi \in \text{Lip}_{d_M}(X, \mathbb{R})$  such that  $\Gamma = \Gamma' = G = \{(x, y) \in X \times X : \varphi(x) - \varphi(y) = d_M(x, y)\}$ . Note that  $\Gamma$  is closed.

The next result shows that the sets  $H(y)$  have the structure of Example 6.1. The convex assumption on the obstacle is fundamental: each transport rays is composed by a straight line, a geodesic on  $M$  where branching structures are allowed and again a straight line.

LEMMA 7.2. *For all  $y \in S$ ,  $H(y)$  has the geometry of Example 6.1: there exists an increasing curve  $\gamma_y : [0, 1] \rightarrow X$  such that*

$$H(y) \times X \cap \Gamma \subset \bigcup_{t \in [0, 1]} K(\gamma_y(t)) \cap \Gamma.$$

*Proof.* Since due to convexity and smoothness of the obstacle, the geodesics of  $d_M$  are smooth and composed by a first straight line, a geodesic of the manifold and a final straight line, a branching structure can appear only on the manifold  $M$ . If  $H(y) \neq R(y)$ , consider the following sets:

$$Z := \bigcap_{z \in H(y) \cap M} G^{-1}(z) \cap M, \quad W := \bigcap_{z \in H(y) \cap M} G(z) \cap M.$$

By  $d_M$ -monotonicity, smoothness and convexity of  $M$ , for all  $z \in H(y) \cap M$  the set  $G^{-1}(z) \cap M$  is always contained in the same geodesic of  $M$ . Using the compactness of  $M$ ,  $Z = \{z\}$  and  $W = \{w\}$  and  $(z, w) \in G$ . Consider the unique increasing geodesic  $\gamma_y \in \gamma_{[z, w]}$  such that  $\gamma_y = G(z) \cap G^{-1}(w)$ . Hence

$$H(y) \times X \cap \Gamma \subset \bigcup_{t \in [0, 1]} K(\gamma_y(t)) \cap \Gamma.$$

□

REMARK 2. From Lemma 7.1 and Lemma 7.2 it follows that, for any transference plan  $\pi$ , the set of chain of transport rays  $H$  satisfies Assumption 1.

Indeed consider  $H(y)$  and the corresponding geodesic  $\gamma_y$  from Lemma 7.2. Then take any sequence  $x_n \in H(y)$  such that  $|x_n - x| \rightarrow 0$  as  $n \rightarrow +\infty$ . Note that there exist  $s_n \in [0, 1]$  and  $t_n \in \mathbb{R}$  such that  $x_n = \gamma_y(s_n) + t_n \nabla \gamma_y(s_n)$ . Possibly passing to subsequences,  $s_n \rightarrow s$ ,  $t_n \rightarrow t$  with  $x = \gamma_y(s) + t \nabla \gamma_y(s)$ . Since  $(\gamma_y(s_n), x_n) \in G$  and  $G$  is closed it follows that  $(\gamma_y(s), x) \in G$ . From  $(y, \gamma_y(s)) \in R$  follows  $x \in H(y)$ . Hence point (a) of Assumption 1 holds true.

Point (b) of Assumption 1 follows directly from the continuity of  $d_N$ .

In the following Lemma we prove that the problem can be reduced to the equivalence classes  $H(y)$ . We use the following notation: the quotient map induced by  $H$  will be denoted by  $f^y$  and the corresponding quotient measure  $f^y_\# \mu$  by  $m_H$ .

LEMMA 7.3. The  $\mu$ -measure of the set of initial points is zero, hence

$$\mu = \int \mu_y m_H(dy).$$

Moreover  $\mu_y$  is continuous for  $m_H$ -a.e.  $y$ .

*Proof.*

*Step 1.* Since  $\mu \ll \mathcal{L}^d$ , it is enough to prove that the set of initial points is  $\mathcal{L}^d$ -negligible and that the disintegration w.r.t.  $H$  of  $\mathcal{L}^d$  restricted to any compact set has continuous conditional probabilities. Indeed if  $\mathcal{L}^d \llcorner K = \int \eta_y m_{\mathcal{L}^d}(dy)$  and  $\mu = \rho \mathcal{L}^d$  then  $m_\mu \ll m_{\mathcal{L}^d}$  and

$$\mu \llcorner K = \int \rho \eta_y m_{\mathcal{L}^d}(dy) = \int \rho \frac{dm_{\mathcal{L}^d}}{dm_\mu} \eta_y m_\mu(dy),$$

where  $m_\mu$  is the quotient measure of  $\mu \llcorner K$ . It follows that the continuity of  $\eta_y$  implies the continuity of conditional probabilities of  $\mu$ . Hence the claim is to prove that  $\mathcal{L}^d$  satisfies Assumption 2.

*Step 2.* Let  $K \subset X$  be any compact set with  $\mathcal{L}^d(K) > 0$ . Possibly intersecting  $K$  with  $B_r(x)$  for some  $x \in \mathbb{R}^d \setminus C$  and  $r > 0$ , we can assume w.l.o.g. that  $K \subset B_\varepsilon(x)$  and  $B_{2\varepsilon}(x) \cap M = \emptyset$ . Since  $d_M \geq d$ ,  $K_t \subset B_{2\varepsilon}(x)$  for all  $t \leq \varepsilon$ . Since  $d_M = |\cdot|$  in  $B_{2\varepsilon}(x)$ , it follows that inside  $B_{2\varepsilon}(x) \times B_{2\varepsilon}(x)$   $d_M$ -cyclically monotonicity is equivalent to  $|\cdot|$ -cyclically monotonicity. It follows that the set  $H \cap G(K) \times G^{-1}(K_\varepsilon)$  is  $|\cdot|$ -cyclically monotone.

*Step 3.* The following is proved in [4]: consider a metric measure space  $(X, d, m)$  with  $d$  non-branching geodesic distance,  $m \in \mathcal{P}(X)$  and assume that  $(X, d, m)$  satisfies  $MCP(K, N)$ , for the definition of  $MCP(K, N)$  we refer to [14]. Let  $\Gamma$  be a  $d$ -cyclically monotone set and consider the evolution of sets induced by  $\Gamma$ , then  $m$  satisfies Assumption 2 w.r.t. this evolution of sets.

Since  $B_{2\varepsilon}(x)$  is a convex set, it follows that  $(B_{2\varepsilon}(x), |\cdot|, \mathcal{L}^d)$  satisfies  $MCP(0, d)$ . Therefore  $\mathcal{L}^d$  satisfies Assumption 2 w.r.t. the evolution of sets induced by  $H \cap G(K) \times G^{-1}(K_\varepsilon)$ . The claim follows. □

Hence we can assume w.l.o.g. that  $\mu(G^{-1}(M)) = \nu(G(M)) = 1$ : if  $H(y)$  do not intersect the obstacle, it is a straight line and the marginal  $\mu_y$  is continuous. Since the existence of an optimal transport map on a straight line with first marginal continuous is a standard fact in optimal transportation, the reduction follows.

Recall the two family of sets introduced in Example 6.1:

$$K_{y,t} := G^{-1}(\gamma_y(t)) \setminus \bigcap_{s < t} G^{-1}(\gamma_y(s)), \quad Q_{y,t} := G(\gamma_y(t)) \setminus \bigcap_{t < s} G(\gamma_y(s)).$$

It follows from Lemma 7.2 and Example 6.1 that

$$\mu_y = \int \mu_{y,t} m_{\mu_y}(dt), \quad \nu_y = \int \nu_{y,t} m_{\nu_y}(dt).$$

with  $\mu_{y,t}(K_{y,t}) = \nu_{y,t}(Q_{y,t}) = 1$ . Moreover using the increasing curve  $\gamma_y$ , we can assume that  $m_{\mu_y} \in \mathcal{P}(M)$ , indeed

$$\mu_y = \int_{[0,1]} \mu_{y,t} m_{\mu_y}(dt) = \int_{\gamma_y([0,1])} \mu_{y,\gamma_y^{-1}(z)} (\gamma_y \# m_{\mu_y})(dz). \quad (7.2)$$

And the same calculation holds true for  $\nu_y$  and  $m_{\nu_y}$ . Therefore in the following

$$\mu_y = \int_M \mu_{y,z} m_{\mu_y}(dz), \quad \nu_y = \int_M \nu_{y,z} m_{\nu_y}(dz) \quad (7.3)$$

with  $\mu_{y,z}(K_{y,\gamma_y^{-1}(z)}) = \nu_{y,z}(Q_{y,\gamma_y^{-1}(z)}) = 1$  and  $m_{\mu_y}(\gamma_y([0,1])) = m_{\nu_y}(\gamma_y([0,1])) = 1$ .

Moreover w.l.o.g. we can assume that  $S = f^y(\mathbb{R}^d) \subset M$ , in particular we can assume that for all  $y \in S$  there exists  $t(y) \in [0,1]$  such that  $y = \gamma_y(t(y))$ .

According to Proposition 6.2, to obtain the existence of an optimal map on  $H(y)$  it is enough to prove that  $m_{\mu_y}$  is continuous and  $\mu_{y,z}$  is continuous for  $m_{\mu_y}$ -a.e.  $z \in M$ . Recall that  $m_{\mu_y}(\gamma_y([0,t])) \geq m_{\nu_y}(\gamma_y([0,t]))$  is a straightforward consequence of the increasing property of  $\gamma_y$ .

REMARK 3. Consider the following  $\mathcal{A}$ -measurable map:

$$G^{-1}(M) \setminus (a(M) \cap M) \ni w \mapsto f^M(w) := \text{Argmin}\{d(z, w) : z \in M \cap G(w)\} \in M.$$

Consider the measure  $m := f_{\#}^M \mu \in \mathcal{P}(M)$ . Observing that  $f^M(H(y)) = \gamma_y([0,1])$ , it follows that the support of  $m$  is partitioned by a  $d_M$ -cyclically monotone equivalence relation:

$$m\left(\bigcup_{y \in S} \gamma_y([0,1])\right) = 1, \quad \bigcup_{y \in S} \gamma_y([0,1]) \times \bigcup_{y \in S} \gamma_y([0,1]) \cap G \text{ is } d_M\text{-cyclically monotone}$$

Moreover  $f^y$  is a quotient map also for this equivalence relation. Note that  $f_{\#}^y m = m_H$ : consider  $I \subset S$

$$\begin{aligned} (f_{\#}^y m)(I) &= m\left(\bigcup_{y \in I} \gamma_y([0,1])\right) = \mu\left(G^{-1}\left(\bigcup_{y \in I} \gamma_y([0,1])\right)\right) \\ &= \mu\left(\bigcup_{y \in I} H(y)\right) = (f_{\#}^y \mu)(I) = m_H(I). \end{aligned}$$

It follows that

$$m = \int_S (f_{\#}^M \mu_y) m_H(dy)$$

and from (7.3)  $f_{\#}^M \mu_y = m_{\mu_y}$ . Hence the final disintegration formula for  $m$  is the following one:

$$m = \int_S m_{\mu_y} m_H(dy). \quad (7.4)$$

PROPOSITION 7.4. The measure  $m$  is absolutely continuous w.r.t. the Hausdorff measure  $\mathcal{H}^{d-1}$  restricted to  $M$ .

Proof. Recall that  $\varphi \in \text{Lip}_{d_M}(\mathbb{R}^d)$  is the potential associated to  $\Gamma$  and consider the following set

$$M_2 := P_1\left(\{(x, y) \in M \times M : |\varphi(x) - \varphi(y)| = d_M(x, y)\} \setminus \{x = y\}\right).$$

*Step 1.* Define the following map:  $M_2 \ni w \mapsto \Xi(w) := \text{Argmin}\{\varphi(w) - \varphi(z) : z \in M_2\}$ . Then the function  $\varphi$  is a potential for the Monge minimization problem on  $M$  with cost the geodesic distance, that coincides with  $d_M$ , with first marginal  $m$  and as second marginal  $\Xi_\# m$ .

It follows from Proposition 15 of [10] that  $\nabla\varphi$  is a Lipschitz function: for all  $x, y \in M_2$

$$|\nabla\varphi(x) - \nabla\varphi(y)| \leq L d_M(x, y).$$

In [10] the Lipschitz constant  $L$  is uniform for  $x, y$  belonging to sets uniformly far from the starting and ending points of the geodesics on  $M$  of the transport set. Since in our setting the geodesics on  $M$  do not intersect,  $L$  is uniform on the whole  $M$ . Moreover note that if  $z = \gamma_y(t)$ , then

$$\nabla\varphi(z) = -\frac{\dot{\gamma}_y(t)}{|\dot{\gamma}_y(t)|}.$$

*Step 2.* For  $t \geq 0$ , define the following map

$$M_2 \ni x \mapsto \psi_t(x) := x + \nabla\varphi(x)t.$$

Possibly restricting  $\psi_t$  to a subset of  $M$  of points coming from transport rays of uniformly positive length, since  $t \mapsto \psi_t(x)$  is a parametrization of the transport ray touching  $M$  in  $x$ , by  $d_M$ -cyclical monotonicity of  $\Gamma$ , we can assume that  $\psi_t$  is injective. Moreover  $\psi_t$  is bi-Lipschitz, provided  $t$  is small enough: indeed

$$|x + \nabla\varphi(x)t - y - \nabla\varphi(y)t| \geq |x - y|(1 - Lt).$$

It follows that

$$M_2 \times [-\delta, \delta] \ni (x, s) \mapsto \psi(x, s) := x + \nabla\varphi(x)(t + s)$$

is bi-Lipschitz and injective provided  $\delta \leq 1/L + t$ . Hence the Jacobian determinant of  $\varphi$ ,  $Jd\varphi$ , is uniformly positive.

*Step 3.* Consider the following set

$$B := \{x \in \mathbb{R}^d : t - \delta \leq d(M, x) \leq t + \delta\} \cap G^{-1}(M)$$

where  $d$  is the euclidean distance. Clearly  $B$  is the range of  $\psi$  and  $\mathcal{L}^d(B) > 0$ . Since  $M$  is a smooth manifold, we can pass to local charts: let  $U_\alpha \subset \mathbb{R}^{d-1}$  be an open set and  $h_\alpha : U_\alpha \rightarrow M$  the corresponding parametrization map. The map

$$U_\alpha \times [-\delta, \delta] \ni (x, s) \mapsto \psi_\alpha(x, s) := \psi(h_\alpha(x), s)$$

is a bi-Lipschitz parametrization of the set  $B_\alpha := B \cap G^{-1}(h_\alpha(U_\alpha))$ .

It follows directly from the Area Formula, see for example [1], that

$$\mathcal{L}^d \llcorner_{B_\alpha} = \psi_{\alpha\#} \left( Jd\psi_\alpha(\mathcal{L}^{d-1} \times dt) \llcorner_{U_\alpha \times [-\delta, \delta]} \right),$$

hence  $f_\#^M \mathcal{L}^d \llcorner_{B_\alpha} \ll \mathcal{H}^{d-1} \llcorner_M$ . Since  $B$  can be covered with a finite number of  $B_\alpha$  and  $\mathcal{L}^d \llcorner_{B_\alpha}$  is equivalent to  $m$ , the claim follows.  $\square$

Recall the following result. Let  $(M, g)$  be a  $n$ -dimensional compact Riemannian manifold, let  $d_M$  be the geodesic distance induced by  $g$  and  $\eta$  the volume measure. Then the disintegration of  $\eta$  w.r.t. any  $d_M$ -cyclically monotone set is strongly consistent and the conditional probabilities are continuous. This result is proved in [4], Theorem 9.5, in the more general setting of metric measure space satisfying the measure contraction property.

**COROLLARY 7.5.** *For  $m_H$ -a.s.  $y \in S$ , the quotient measure  $m_{\mu_y}$  is continuous.*

*Proof.* We have proved in Remark 3 that the measures  $m_{\mu_y}$  are the conditional probabilities of the disintegration of  $m$  w.r.t. the equivalence relation given by the membership to geodesics  $\gamma_y$  and  $m_H$  is the corresponding quotient measure. Hence the claim follows directly from Theorem 9.5 of [4] and Proposition 7.4.  $\square$

PROPOSITION 7.6. *For  $m_H$ -a.e.  $y \in S$ , the measures  $\mu_{y,z}$  are continuous for  $m_{\mu_y}$ -a.e.  $z \in M$ .*

*Proof.* Recall that  $f_{\#}^M \mu = m$ .

*Step 1.* The measure  $\mu$  can be disintegrated w.r.t. the partition given by the family of pre-images of the  $\mathcal{A}$ -measurable map  $f^M: \{(f^M)^{-1}(p)\}_{p \in f^M(\mathbb{R}^d)}$ . Clearly  $f^M$  is a possible quotient map, hence

$$\mu = \int \mu_z m(dz), \quad (7.5)$$

The set  $G^{-1}(M) \setminus a(M) \times G^{-1}(M) \setminus a(M) \cap G$  is  $|\cdot|$ -cyclically monotone and  $\mu \ll \mathcal{L}^d$ , hence it follows that for  $m$ -a.e.  $z \in f^M(\mathbb{R}^d)$ ,  $\mu_z$  is continuous.

*Step 2.* From Lemma 7.3  $\mu = \int \mu_y m_H(dy)$ , therefore

$$m = f_{\#}^M \mu = \int (f_{\#}^M \mu_y) m_H(dy),$$

hence using (7.5) and the uniqueness of the disintegration

$$\mu = \int \left( \int \mu_z (f_{\#}^M \mu_y)(dz) \right) m_H(dy), \quad \mu_y = \int \mu_z (f_{\#}^M \mu_y)(dz),$$

where the last equality holds true for  $m_H$ -a.e.  $y \in S$ . Hence for  $m_H$ -a.e.  $y \in S$  the measures  $\mu_{y,z}$  are continuous for  $m_{\mu_y}$ -a.e.  $z \in M$ .  $\square$

Finally we can prove the existence of an optimal map for the Monge minimization problem with obstacle.

THEOREM 7.7. *There exists a solution for the Monge minimization problem with cost  $d_M$  and marginal  $\mu, \nu$  with  $\mu \ll \mathcal{L}^d$ .*

*Proof.* From Lemma 7.3 it follows that  $\mu$  can be disintegrated w.r.t. the equivalence relation  $H$ . From Theorem 6.1 it follows that to prove the claim it is enough to prove the existence of an optimal map on each equivalence class  $H(y)$ . Hence we restrict the analysis to the classes  $H(y)$  such that  $H(y) \neq R(y)$  and for them we proved in Lemma 7.2 that Assumption 3 holds true. In Proposition 7.4, Corollary 7.5 and Proposition 7.6 we proved that for  $m_H$ -a.e.  $y \in S$  the measures  $m_{\mu_y}$  and  $\mu_{y,z}$  verify the hypothesis of Proposition 6.2. Therefore the claim follows.  $\square$

#### ACKNOWLEDGEMENTS

I would like to express my gratitude to Stefano Bianchini for his support during the preparation of this paper. I wish to thank Luigi De Pascale for drawing my attention on the obstacle problem.

#### Appendix A. Notation.

$P_{i_1 \dots i_I}$	projection of $x \in \Pi_{k=1, \dots, K} X_k$ into its $(i_1, \dots, i_I)$ coordinates, keeping order
$\mathcal{P}(X)$ or $\mathcal{P}(X, \Omega)$	probability measures on a measurable space $(X, \Omega)$
$\mathcal{M}(X)$ or $\mathcal{M}(X, \Omega)$	signed measures on a measurable space $(X, \Omega)$
$f _A$	the restriction of the function $f$ to $A$
$\mu _A$	the restriction of the measure $\mu$ to the $\sigma$ -algebra $A \cap \Sigma$
$\mathcal{L}^d$	Lebesgue measure on $\mathbb{R}^d$
$\mathcal{H}^k$	$k$ -dimensional Hausdorff measure
$\Pi(\mu_1, \dots, \mu_I)$	$\pi \in \mathcal{P}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$ with marginals $(P_i)_{\#} \pi = \mu_i \in \mathcal{P}(X_i)$
$\mathcal{I}(\pi)$	cost functional (2.5)
$c$	cost function : $X \times Y \mapsto [0, +\infty]$

$\mathcal{I}$	transportation cost (2.5)
$\phi^c$	$c$ -transform of a function $\phi$ (2.6)
$\partial^c \varphi$	$d$ -subdifferential of $\varphi$ (2.7)
$\Phi_c$	subset of $L^1(\mu) \times L^1(\nu)$ defined in (2.8)
$J(\phi, \psi)$	functional defined in (2.9)
$C_b$ or $C_b(X, \mathbb{R})$	continuous bounded functions on a topological space $X$
$(X, d)$	Polish space
$(X, d_L)$	non-branching geodesic separable metric space
$D_N(x)$	the set $\{y : d_N(x, y) < +\infty\}$
$L(\gamma)$	length of the Lipschitz curve $\gamma$ , Definition 2.8
$B_r(x)$	open ball of center $x$ and radius $r$ in $(X, d)$
$B_{r,L}(x)$	open ball of center $x$ and radius $r$ in $(X, d_L)$
$\mathcal{K}(X)$	space of compact subsets of $X$
$d_H(A, B)$	Hausdorff distance of $A, B$ w.r.t. the distance $d$
$A_x, A^y$	$x, y$ section of $A \subset X \times Y$ (2.3)
$\mathcal{B}, \mathcal{B}(X)$	Borel $\sigma$ -algebra of $X$ Polish
$\Sigma_1^1, \Sigma_1^1(X)$	the pointclass of analytic subsets of Polish space $X$ , i.e. projection of Borel sets
$\Pi_1^1$	the pointclass of coanalytic sets, i.e. complementary of $\Sigma_1^1$
$\Sigma_n^1, \Pi_n^1$	the pointclass of projections of $\Pi_{n-1}^1$ -sets, its complementary
$\Delta_n^1$	the ambiguous class $\Sigma_n^1 \cap \Pi_n^1$
$\mathcal{A}$	$\sigma$ -algebra generated by $\Sigma_1^1$
$\mathcal{A}$ -function	$f : X \rightarrow \mathbb{R}$ such that $f^{-1}((t, +\infty])$ belongs to $\mathcal{A}$
$h_{\#}\mu$	push forward of the measure $\mu$ through $h$ , $h_{\#}\mu(A) = \mu(h^{-1}(A))$
$\text{graph}(F)$	graph of a multifunction $F$ (2.1)
$F^{-1}$	inverse image of multifunction $F$ (2.2)
$F_x, F^y$	sections of the multifunction $F$ (2.3)
$\text{Lip}_1(X)$	Lipschitz functions with Lipschitz constant 1
$\Gamma'$	transport set (3.1)
$G, G^{-1}$	outgoing, incoming transport ray, Definition 3.2
$R$	set of transport rays (3.3)
$\mathcal{T}, \mathcal{T}_e$	transport sets (3.4)
$a, b : \mathcal{T}_e \rightarrow \mathcal{T}_e$	endpoint maps (3.7)
$\mathcal{Z}_{m,e}, \mathcal{Z}_m$	partition of the transport set $\Gamma$ (4.1), (4.2)
$S$	cross-section of $R_{\mathcal{L}\mathcal{T} \times \mathcal{T}}$
$A_t$	evolution of $A \subset \mathcal{Z}_{k,i,j}$ along geodesics (5.1)

## REFERENCES

- [1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs, 2000.
- [2] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.
- [3] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin.(N.S.)*, 4(4):353–454, 2009.
- [4] S. Bianchini and F. Cavalletti. The Monge problem for distance cost in geodesic spaces. preprint SISSA, 2010.
- [5] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*. Graduate studies in mathematics. AMS, 2001.
- [6] L. Caffarelli, M. Feldman, and R.J. McCann. Constructing optimal maps for Monge’s transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2002.
- [7] P. Cardaliaguet and C. Jimenez. Optimal transport with convex obstacle. *preprint*, 2010.
- [8] T. Champion and L. De Pascale. The Monge problem in  $\mathbb{R}^d$ . *Duke Math. J.*, 157(3):551–572, 2011.
- [9] L.C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.



- [10] M. Feldman and R. McCann. Monge's transport problem on a Riemannian manifold. *Trans. Amer. Math. Soc.*, 354:1667–1697, 2002.
- [11] D. H. Fremlin. *Measure Theory*, volume 4. Torres Fremlin, 2002.
- [12] D.G. Larman. A compact set of disjoint line segments in  $\mathbb{R}^3$  whose end set has positive measure. *Mathematika*, 18:112–125, 1971.
- [13] A. M. Srivastava. *A course on Borel sets*. Springer, 1998.
- [14] K.T. Sturm. On the geometry of metric measure spaces.II. *Acta Math.*, 196(1):133–177, 2006.
- [15] V.N. Sudakov. Geometric problems in the theory of dimensional distributions. *Proc. Steklov Inst. Math.*, 141:1–178, 1979.
- [16] N. Trudinger and X.J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, 13:19–31, 2001.
- [17] C. Villani. *Topics in optimal transportation*. Graduate studies in mathematics. AMS, 2003.
- [18] C. Villani. *Optimal transport, old and new*. Springer, 2008.